Teachers' Fixed-Ability Practices: Measurement, Antecedents, and Implications for Students' Motivational Beliefs and Achievement

Pia Kreijkes

Christ's College, University of Cambridge

This thesis is submitted for the degree of Doctor of Philosophy

Faculty of Education

September 2020

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. The text does not exceed the prescribed word limit as given by the Degree Committee of the Faculty of Education.

The following is an abstract of the thesis:
"Teachers' fixed-ability practices: Measurement, antecedents, and implications for students' motivational beliefs and achievement"
authored by Pia Kreijkes

Mindsets (i.e., implicit beliefs about whether ability can be changed) are important for students' learning, but relatively little is known about how they develop. Considering the gaps in the literature, the overarching purpose of the present study was to understand the role of teaching practices in the development of students' mindsets and associated achievement-related motivational beliefs. This study set out to first develop a conceptualisation of, and a teacher and student questionnaire for, the assessment of fixed-ability practices (FAPs) – teaching practices that are based on the notion that ability is fixed. It aimed to examine the dimensionality of student-perceived FAPs (RQ 1), and whether teachers' own motivational beliefs (mindset, teaching goal orientation, attribution for students' achievement, and perceived student ability/expectation) predict their engagement in student-perceived and/or teacher-reported FAPs (RQ 2). Moreover, it aimed to examine whether student-perceived and/or teacher-reported FAPs predict students' motivational beliefs (mindset, achievement goal orientation, achievement attribution, and ability self-concept) as well as achievement (RQ 3). To gain a greater understanding of why such relationships might exist, this study also explored how teachers justify their engagement in FAPs, and whether these justifications are related to their motivational beliefs (RQ 4). It also explored how students perceive FAPs, in particular, what messages they draw from these for their motivational beliefs (RQ 5). The research questions were examined in a mathematics context because maths is a subject in which fixed mindsets are particularly prevalent.

A longitudinal mixed-methods design was used. Participants were 927 Year 7 students (aged 11-12 years) from 37 classes and 31 of their maths teachers in seven English secondary schools. Teachers' and students' motivational beliefs in maths were assessed with self-report questionnaires at the beginning of the first year of secondary school (Time 1), and students' motivational beliefs were again assessed at the end of the school year (Time 3). During the middle of the year, FAPs were assessed with teacher and student questionnaires (Time 2). Additionally, classroom observations (Time 2) were conducted during maths lessons of four selected teachers. This was followed by video-stimulated recall interviews with these teachers and some of their students to reflect on FAPs. Lastly, students' maths grades for the end of Year 6 and Year 7 were obtained from school records.

Two-level factor analyses showed that student-perceived FAPs comprise five dimensions at

the individual level: unsophisticated task differentiation, public evaluation, little promotion of self-regulation, differential value, and little promotion of risk-taking. There was no overarching FAPs dimension that represented all these practices well. At the classroom level, there was one uniform dimension. Corresponding FAPs scales for teachers had inadequate psychometric properties and were thus not used in subsequent analyses. Correlations between teachers' motivational beliefs and students' shared perceptions of FAPs were all non-significant. Hence, initially planned regression analyses were not performed to examine these relationships further. Longitudinal effects of student-perceived FAPs on students' motivational beliefs and achievement were examined using Bayesian cross-classified multiple membership models, accounting for students belonging to multiple classrooms and teachers during the school year. Results indicated that students' individual perceptions of teachers' differential value led to stronger fixed mindsets and a reduced self-concept. Moreover, perceived little promotion of risk-taking led to more performance-approach goals. In addition, there were several notable trends. At the classroom and teacher level, students' shared perceptions of FAPs led to an increase in performance-approach goals.

Thematic analysis of teacher interviews showed that teachers provided multiple justifications for their practices involving beliefs, goals for students, practical considerations, and school policies. Engagement in more adaptive practices was usually explained by adaptive beliefs and values, while maladaptive practices were explained by maladaptive beliefs and values. Yet, teachers rarely referred to their mindset or goal orientation, and relationships between beliefs and practices were at times inconsistent. As expected, discussions with students indicated that they perceived FAPs negatively and that FAPs seemed to foster maladaptive motivational beliefs. Moreover, students' initial beliefs seemed to influence how they perceived some practices. The qualitative and quantitative results were mostly aligned, but student interviews revealed additional potential influences of FAPs on motivational beliefs.

The current study has made important theoretical, empirical, and methodological contributions to the field of educational psychology and, in particular, motivation. In addition, the findings have valuable practical implications as they may help teachers and other educational stakeholders understand how to avoid unwittingly setting up students to develop maladaptive motivational beliefs.

Acknowledgements

I could not have written this PhD thesis without the support, advice, and encouragement of many people along the way. First, I would like to thank my supervisor Ros McLellan for being a truly brilliant mentor. Your inspiration and invaluable support have guided this thesis and helped me develop as a researcher. I am deeply grateful.

Many thanks to my adviser Jan Vermunt for his insightful advice regarding the study design as well as Steven Watson for his constructive feedback following the upgrade viva. I would also like to thank my colleague and dear friend Junlin Yu for his help throughout the years and our constant exchange of ideas. My thanks also go to Sonia Ilie for her statistical advice, Charlotte Allen for coding part of my interview data, and Rebecca Gordon for proof reading part of this thesis. I would also like to thank my PhD examiners Andreas Stylianides and Lars-Erik Malmberg for a very interesting discussion about this thesis.

I am grateful to the German Academic Scholarship Foundation for providing generous funding for this PhD (and previous degrees) and for their stimulating development programme. Similarly, I thank Christ's College and the Faculty of Education for providing conference and fieldwork funding and, more importantly, their welcoming community and the various social and academic events, where I met fantastic people and formed wonderful friendships.

Invaluable was the help of my research participants. Many thanks to all the schools, teachers, and students who opened their classrooms, devoted their valuable time, and shared their thoughts for this study.

Last but certainly not least, I would like to thank my family and loved ones. Many thanks to all my dear friends, in particular those of the "Cam Fam," who have made my time in Cambridge so memorable. A heartfelt thank you to my family, especially my parents, for supporting me throughout everything. I will forever be grateful. I cannot express how sad I am that my father is no longer with me to see the end of this PhD journey. He would have been so proud. Many (many!) thanks to my partner Alessandro for being on my side every step of the way. Your love and unwavering confidence were the greatest encouragement. And, finally, I am excited to thank a very special little person who has been my constant companion over the last months. Feeling you move and thinking about your arrival have been a great motivational force and have given me so much perspective.

Table of Contents

Abstract **iii**
Acknowledgements **v**
List of Tables **xiii**
List of Figures **xvi**

Chapter 1 Introduction **1**
1.1 Background for this thesis 1
1.2 The present study 3

Chapter 2 Literature Review **5**
2.1 The meaning system of mindset and related motivational beliefs 6
2.1.1 Achievement goal orientation 7
2.1.2 Achievement attribution 8
2.1.3 Ability self-concept, perceived ability, and expectations 9
2.2 A conceptualisation of fixed-ability practices 11
2.2.1 Tasks 14
2.2.1.1 Teachers' motivational beliefs and tasks 15
2.2.1.2 Tasks and students' motivational beliefs and achievement 16
2.2.1.3 Brief summary on tasks and mindset 17
2.2.2 Grouping 17
2.2.2.1 Teachers' motivational beliefs and grouping 18
2.2.2.2 Grouping and students' motivational beliefs and achievement 18
2.2.2.3 Brief summary on grouping and mindset 19
2.2.3 Evaluation 20
2.2.3.1 Teachers' motivational beliefs and evaluation 20
2.2.3.2 Evaluation and students' motivational beliefs and achievement 22
2.2.3.3 Brief summary on evaluation and mindset 24
2.2.4 Little autonomy 24
2.2.4.1 Teachers' motivational beliefs and little autonomy 25
2.2.4.2 Little autonomy and students' motivational beliefs and achievement 25
2.2.4.3 Brief summary on little autonomy and mindset 26
2.2.5 Differential value 26
2.2.5.1 Teachers' motivational beliefs and differential value 27
2.2.5.2 Differential value and students' motivational beliefs and achievement 27
2.2.5.3 Brief summary on differential value and mindset 28
2.2.6 Little promotion of risk-taking 28
2.2.6.1 Teachers' motivational beliefs and little promotion of risk-taking 28
2.2.6.2 Little promotion of risk-taking and students' motivational beliefs and achievement 29
2.2.6.3 Brief summary on little promotion of risk-taking and mindset 30
2.3 Objective and research questions 31
2.3.1 Need for study and objective 31
2.3.2 Research questions 34

Chapter 3 Research Methodology **35**
3.1 Research design and theoretical perspective 35
3.2 Research methods 38
3.2.1 Assessing motivational beliefs 38

3.2.2 Assessing fixed-ability practices 39
3.2.3 Exploring why teachers use fixed-ability practices and how these affect students' motivational beliefs 40
3.3 Sampling, setting, and participants 40
3.3.1 Sampling 40
3.3.1.1 The overall sample 41
3.3.1.2 The sample for the mixed-methods strand 42
3.3.2 Setting 42
3.3.3 Participants 44
3.3.3.1 Participants of the overall sample 44
3.3.3.2 Participants of the mixed-methods strand 45
3.4 Quantitative research instruments 45
3.4.1 Motivational beliefs questionnaires 47
3.4.1.1 Teacher motivational beliefs questionnaire 47
3.4.1.1.1 Teacher mindset 47
3.4.1.1.2 Teacher goal orientation 50
3.4.1.1.3 Teacher achievement attribution 50
3.4.1.1.4 Teacher perceived ability and expectations 50
3.4.1.2 Student motivational beliefs questionnaire 50
3.4.1.2.1 Student mindset 50
3.4.1.2.2 Student goal orientation 51
3.4.1.2.3 Student achievement attribution 51
3.4.1.2.4 Student self-concept 51
3.4.2 Fixed-ability practices questionnaire 52
3.5 Qualitative research instruments 56
3.5.1 Interview schedule 56
3.6 Maths achievement 57
3.7 Procedure 57
3.7.1 Pilot studies 57
3.7.2 Procedures of the main study 58
3.7.2.1 Time 1: Teacher and student motivational beliefs questionnaires 59
3.7.2.2 Time 2: FAPs questionnaires, observations, and interviews 59
3.7.2.2.1 Procedures for the mixed-methods strand 59
3.7.2.2.2 Procedures for the entire sample 60
3.7.2.3 Time 3: Student motivational beliefs questionnaire 61
3.8 Ethical considerations 61

Chapter 4 General Quantitative Analytical Strategy and Analysis Approach for FAPs Dimensionality (RQ 1) 63
4.1 The general quantitative analytical strategy 63
4.1.1 Overview of the general quantitative analytical strategy 63
4.1.2 Data structure and multilevel modelling 64
4.1.3 Missing data and careless responses 65
4.2 Analyses for research question 1 68
4.2.1 Analyses overview 68
4.2.2 Model evaluation criteria 70
4.2.3 Exploratory factor analysis in the exploratory sample 70
4.2.4 Confirmatory ESEM in the confirmatory sample 71
4.2.4.1 First-order confirmatory ESEM 71
4.2.4.2 Hierarchical ESEM 72
4.2.4.3 Bifactor ESEM 72
4.2.4.4 Model comparison 73

4.2.5 Two-level factor analysis in the entire sample 73

Chapter 5 Results for FAPs Dimensionality (RQ 1) 74
5.1 Results of ESEM 74
5.2 Results of confirmatory ESEM 77
5.2.1 First-order confirmatory ESEM 77
5.2.2 Hierarchical ESEM 78
5.2.3 Bifactor ESEM 79
5.2.4 Model comparison 79
5.3 Results of two-level ESEM 79
5.4 Psychometric properties of student-perceived fixed-ability practices scales 82
5.5 Teacher-reported fixed-ability practices 84
5.6 Brief summary on RQ 1 85

Chapter 6 Analysis Approach for Quantitative Relationships (RQs 2 and 3) 86
6.1 Analyses overview 86
6.2 Data preparation and preliminary analyses 87
6.2.1 Composite scores 88
6.2.2 Standardised achievement scores 88
6.2.3 Weights for group membership 88
6.2.4 Longitudinal invariance testing 89
6.3 Examining the effect of student-perceived fixed-ability practices on students' motivational beliefs and achievement 90
6.3.1 Overview of the cross-classified multiple membership models 90
6.3.2 Model fitting choices 91
6.3.2.1 Distribution family 91
6.3.2.2 Inclusion of random effects 92
6.3.2.3 Centring of predictors 93
6.3.2.4 Selection of priors 94
6.3.2.5 Model estimation 95
6.3.3 Model checking 96
6.3.3.1 Model convergence 96
6.3.3.2 Model fit: Posterior predictive checks 97
6.3.4 Interpretation of regression results 97

Chapter 7 Results for Quantitative Relationships (RQs 2 and 3) 98
7.1 Preliminary results 98
7.1.1 Descriptive statistics and correlations 98
7.1.2 Longitudinal invariance 102
7.1.3 Model convergence 103
7.1.4 Model fit: Posterior predictive checks 103
7.2 Longitudinal effects of student-perceived FAPs on students' motivational beliefs and achievement 105
7.2.1 Unconditional models 105
7.2.2 Conditional models 108
7.2.2.1 Student-level effects 111
7.2.2.2 Classroom/teacher-level effects 114
7.3 Brief summary on RQs 2 and 3 115

Chapter 8 Qualitative Analysis Approach for Interviews (RQs 4 and 5) 116
8.1 Overview of the thematic analysis 116
8.2 Steps of the thematic analysis 116

8.2.1 Generation of codes 118
8.2.1.1 Intercoder reliability 120
8.2.2 Identification and definition of themes 121

Chapter 9 Interview Themes and Findings of Teacher Interviews (RQ 4) 124
9.1 The participants and setting 124
9.2 The themes: Fixed-ability practices and motivational beliefs 125
9.2.1 Fixed-ability practices 126
9.2.1.1 Tasks 128
9.2.1.2 Grouping 129
9.2.1.3 Evaluation 129
9.2.1.4 Little autonomy 130
9.2.1.5 Differential value 130
9.2.1.6 Little promotion of risk-taking 130
9.2.2 Motivational beliefs 131
9.2.2.1 Mindset 135
9.2.2.1.1 Teacher mindset 135
9.2.2.1.2 Student mindset 135
9.2.2.2 Achievement goal orientation 135
9.2.2.2.1 Teacher goal orientation 135
9.2.2.2.2 Student goal orientation 136
9.2.2.3 Attribution 136
9.2.2.3.1 Teacher attribution 136
9.2.2.3.2 Student attribution 136
9.2.2.4 Student self-concept 136
9.2.2.5 Teachers' perceived ability and expectations for students 137
9.2.2.6 Nature of maths 137
9.2.2.7 Student emotion 137
9.3 Teachers' justifications for engaging in fixed-ability practices 138
9.3.1 Teachers' motivational beliefs and justifications for tasks 138
9.3.1.1 Unsophisticated task differentiation 138
9.3.1.2 Unidimensional tasks 141
9.3.2 Teachers' motivational beliefs and justifications for grouping 142
9.3.2.1 Group and pair work 142
9.3.2.2 Seating and ability grouping 144
9.3.2.3 Stability 144
9.3.3 Teachers' motivational beliefs and justifications for evaluation 144
9.3.3.1 Public evaluation 144
9.3.3.2 Performance over process emphasis 147
9.3.4 Teachers' motivational beliefs and justifications for little autonomy 150
9.3.4.1 Little choice 150
9.3.4.2 Little promotion of self-regulation 151
9.3.5 Teachers' motivational beliefs and justifications for differential value 152
9.3.6 Teachers' motivational beliefs and justifications for little promotion of risk-taking 153
9.3.6.1 Little encouragement of help-seeking 153
9.3.6.2 Unconstructive handling of mistakes 153
9.3.7 Summary on teachers' justifications for their teaching practices 154

Chapter 10 Findings of Student Interviews (RQ 5) 157
10.1 Students' perceptions of fixed-ability practices in relation to their motivational beliefs 157

10.1.1 Tasks and students' motivational beliefs 157
10.1.1.1 Unsophisticated task differentiation 157
10.1.1.2 Unidimensional tasks 159
10.1.2 Grouping and students' motivational beliefs 160
10.1.2.1 Group and pair work 160
10.1.2.2 Seating and ability grouping 160
10.1.2.3 Stability 161
10.1.3 Evaluation and students' motivational beliefs 162
10.1.3.1 Public evaluation 162
10.1.3.2 Performance over process emphasis 164
10.1.4 Autonomy and students' motivational beliefs 166
10.1.4.1 Little choice 166
10.1.4.2 Little promotion of self-regulation 166
10.1.5 Differential value and students' motivational beliefs 167
10.1.6 Little promotion of risk-taking and students' motivational beliefs 168
10.1.6.1 Little encouragement of help-seeking 168
10.1.6.2 Unconstructive handling of mistakes 169
10.1.7 Summary on students' perceptions of FAPs 170

Chapter 11 Discussion **173**
11.1 What are fixed-ability practices? 174
11.1.1 Brief summary on what fixed-ability practices are 178
11.2 Why do teachers engage in fixed-ability practices? 179
11.2.1 Tasks 181
11.2.2 Grouping 183
11.2.3 Evaluation 186
11.2.4 Little autonomy 188
11.2.5 Differential value 189
11.2.6 Little promotion of risk-taking 190
11.2.7 Brief summary on why teachers engage in fixed-ability practices 192
11.3 How do fixed-ability practices affect students' motivational beliefs and achievement? 192
11.3.1 Tasks 193
11.3.2 Grouping 194
11.3.3 Evaluation 196
11.3.4 Little autonomy 198
11.3.5 Differential value 199
11.3.6 Little promotion of risk-taking 200
11.3.7 Absence of effects on achievement 202
11.3.8 Classroom level effects 203
11.3.9 Brief summary on how fixed-ability practices affect students' motivational beliefs and achievement 204

Chapter 12 Conclusions **205**
12.1 Contributions and implications of the study 205
12.1.1 Conceptualisation and measurement of fixed-ability practices 205
12.1.2 Other methodological contributions 207
12.1.3 Teachers' motivational beliefs as antecedents of fixed-ability practices 208
12.1.4 Implications of fixed-ability practices for students' motivational beliefs 208
12.1.5 Implications for educational stakeholders 209
12.2 Limitations and future directions for research 213
12.2.1 Quantitative and qualitative limitations 213

12.2.2 New directions for future research 216
12.3 Concluding remarks 220

References 221
Appendix A Questionnaires Administered in the Main Study 251
A.1 Teacher motivational beliefs questionnaire 251
A.1.1 Demographic questions 253
A.2 Student motivational beliefs questionnaire 254
A.2.1 Demographic questions 256
A.3 Fixed-ability practices questionnaires for teachers and students 257
Appendix B Example Interview Schedules 262
B.1 Example teacher interview schedule 262
B.2 Example student group interview schedule 266
Appendix C Pilot Studies 271
C.1 Pilot study 1 271
C.1.1 Participants and procedure 271
C.1.2 Analysis, results, and discussion 271
C.1.2.1 Quantitative instruments 271
C.1.2.1.1 Motivational beliefs questionnaire 272
C.1.2.1.2 Fixed-ability practices questionnaire 273
C.1.2.2 Qualitative instruments 274
C.1.2.2.1 Observations 274
C.1.2.2.2 Interviews 281
C.1.3 Conclusions 282
C.2 Pilot study 2 282
C.2.1 Participants and procedure 282
C.2.2 Analysis, results, and discussion 282
C.2.3 Conclusions 283
Appendix D Information and Informed Consent Forms for Parents/Guardians, Teachers, and Students 290
D.1 Information and opt-out form for parents/guardians 290
D.2 Information for parents/guardians for the mixed-methods sample 292
D.3 Information and written informed assent form for students 294
D.4 Information and written informed consent form for teachers 296
Appendix E Example Notes of Classroom Observation 299
Appendix F Instructions for Questionnaire Administration 301
Appendix G Results for Exploratory Structural Equation Modelling 302
Appendix H Technical Details for Regression Models 306
H.1 Irregularities in assigning weights for group membership 306
H.2 Formulae for regression models 307
H.3 Priors for Bayesian regression analyses 311
Appendix I Results for the Longitudinal Measurement Invariance Testing 313
Appendix J Additional Results for the CC and CCMM Models 315
J.1 Posterior predictive checks (PPC) for the final CC regression model 315
J.2 Posterior predictive checks (PPC) for the final CCMM regression model 318
J.3 The unconditional and conditional CC and CCMM models with brms default priors only 321
Appendix K Example Summary of Student Interview Resulting from Data Familiarisation 324
Appendix L Steps in the Coding Process and Prominent Changes in Codes 326
Appendix M Coding Framework for Fixed-Ability Practices and Motivational Beliefs 329

Appendix N Examples for a Teacher and Student Matrix **339**
Appendix O Themes in Teacher and Student Interviews **345**
O.1 Fixed-ability practices 345
O.1.1 Tasks 345
O.1.1.1 Unsophisticated task differentiation 345
O.1.1.2 Unidimensional tasks 345
O.1.2 Grouping 346
O.1.2.1 Group and pair work 346
O.1.2.2 Seating and ability grouping 346
O.1.2.3 Stability 348
O.1.3 Evaluation 348
O.1.3.1 Public evaluation 348
O.1.3.2 Performance over process emphasis 350
O.1.4 Little autonomy 353
O.1.4.1 Little choice 353
O.1.4.2 Little promotion of self-regulation 354
O.1.5 Differential value 355
O.1.6 Little promotion of risk-taking 356
O.1.6.1 Little encouragement of help-seeking 356
O.1.6.2 Unconstructive handling of mistakes 357
O.2 Motivational beliefs 359
O.2.1 Mindset 359
O.2.1.1 Teacher mindset 359
O.2.1.2 Student mindset 360
O.2.2 Achievement goal orientation 361
O.2.2.1 Teacher goal orientation 361
O.2.2.2 Student goal orientation 361
O.2.3 Attribution 363
O.2.3.1 Teacher attribution 363
O.2.3.2 Student attribution 364
O.2.4 Student self-concept 365
O.2.5 Teachers’ perceived ability and expectations for students 365
O.2.6 Nature of maths 367
O.2.7 Student emotion 367

List of Tables

Table 2.1 Dimensions and Characteristics of Fixed-Ability Practices 13
Table 3.1 Overview of the Research Questions 36
Table 3.2 School Characteristics 43
Table 3.3 Demographic Information of the Teacher Sample 44
Table 3.4 Characteristics of the Mixed-Methods Teachers 45
Table 3.5 Overview of the Teacher and Student Questionnaires 46
Table 3.6 Example Items and Prior Psychometric Properties of Teacher Motivational Beliefs Questionnaire 48
Table 3.7 Example Items and Prior Psychometric Properties of Student Motivational Beliefs Questionnaire 49
Table 3.8 Proposed Scales and Example Items of Fixed-Ability Practices Questionnaire for Teachers and Students 55
Table 5.1 Results for the Exploratory Factor Analysis of Fixed-Ability Practices for Five Factors: Standardised Loadings and Factor Correlations 77
Table 5.2 Final ESEM Model Fit Indices for Split-Half Sample 2 78
Table 5.3 Results from the Two-Level Exploratory Factor Analyses of Fixed-Ability Practices Items: Model Fit of Models with Five Within-Level Factors and Varying Numbers of Between-Level Factors 80
Table 5.4 Results from Two-Level Exploratory Factor Analysis of Fixed-Ability Practices Items: Within- and Between-Level Loadings for a Model with Five Student-Level Factors and One Classroom-Level Factor 81
Table 5.5 Scales and Items of the Final Fixed-Ability Practices Questionnaire for Students 83
Table 5.6 Descriptive Statistics for Student-Reported Fixed-Ability Practices 84
Table 5.7 Descriptive Statistics for Teacher-Reported Fixed-Ability Practices 84
Table 7.1 Scale Descriptive Statistics 100
Table 7.2 Pearson's Correlations Amongst Students' Motivational Beliefs and Student-Perceived Fixed-Ability Practices 101
Table 7.3 Pearson's Correlations Amongst Teachers' Motivational Beliefs 102
Table 7.4 Pearson's Correlations Between Teachers' Motivational Beliefs and Student-Perceived FAPs 102
Table 7.5 Summary of Model Fit Statistics for Testing Measurement Invariance Across Time 104
Table 7.6 Results of the Two-Level Cross-Classified Random Intercept Model Predicting Students' Motivational Beliefs and Achievement Using the Single-Teacher Sample 106
Table 7.7 Results of the Two-Level Cross-Classified Multiple Membership Random Intercept Model Predicting Students' Motivational Beliefs and Achievement Using the Entire Sample 107
Table 7.8 Results of the Two-Level Cross-Classified Random Effects Model Predicting Students' Motivational Beliefs and Achievement Using the Single-Teacher Sample 109
Table 7.9 Results of the Two-Level Cross-Classified Multiple Membership Random Effects Model Predicting Students' Motivational Beliefs and Achievement Using the Entire Sample 110
Table 8.1 Steps of Thematic Analysis 117
Table 8.2 Example of a Coded Extract 119
Table 9.1 Teachers' Fixed-Ability Practices Reported in Teacher and Student Interviews 128
Table 9.2 Teachers' Motivational Beliefs as Reported in Interviews 133

Table 9.3 Students' Motivational Beliefs as Reported in Interviews 134
Table 9.4 Teachers' Justification for Their Engagement in Fixed-Ability Practices Revealed During Interviews 139
Table 10.1 Potential Influences of Fixed-Ability Practices on Students' Motivational Beliefs as Revealed During Student Interviews 158
Table A.1 Teacher Motivational Beliefs Questionnaire 251
Table A.2 Student Motivational Beliefs Questionnaire 254
Table A.3 Fixed-Ability Practices Questionnaire for Teachers and Students 257
Table C.1 Descriptive Statistics and Internal Reliability for Teacher Motivational Beliefs Scales 272
Table C.2 Descriptive Statistics and Internal Reliability for Student Motivational Beliefs Scales 273
Table C.3 Descriptive Statistics and Internal Reliability for Teacher Fixed-Ability Practices Scales 273
Table C.4 Descriptive Statistics and Internal Reliability for Student Fixed-Ability Practices Scales 274
Table C.5 Overview of the Modifications of Student Fixed-Ability Practices Items After Pilot Study 1 275
Table C.6 Descriptive Statistics and Internal Reliability for Student Fixed-Ability Practices Scales in Pilot Study 2 283
Table C.7 Overview of the Modifications of Student Fixed-Ability Practices Items After Pilot Study 2 284
Table E.1 Example Notes of Classroom Observation 299
Table G.1 Results of the First-Order Confirmatory ESEM: Factor Loadings and Standardised Factor Correlations 302
Table G.2 Results of the Hierarchical Confirmatory ESEM: Factor Loadings and Standardised Higher-Order Uniform Factor Loadings 303
Table G.3 Results of the Bifactor ESEM: Factor Loadings 304
Table G.4 Results of the First-Order Confirmatory ESEM with the Entire Sample: Factor Loadings and Standardised Factor Correlations 305
Table H.1 Priors Used in the Bayesian Regression Models 311
Table I.1 Results of the Unconstrained Configural ESEM for Motivational Beliefs: Standardised Loadings and Factor Correlations for Time 1 313
Table I.2 Results of the Unconstrained Configural ESEM for Motivational Beliefs: Standardised Loadings and Factor Correlations for Time 3 314
Table I.3 Standardised Factor Correlations for Motivational Beliefs Across Time Points 314
Table J.1 Results of the Two-Level Cross-Classified Random Intercept Model Predicting Students' Beliefs and Achievement Using the Single-Teacher Sample 321
Table J.2 Results of the Two-Level Cross-Classified Multiple-Membership Random Intercept Model Predicting Students' Beliefs and Achievement Using the Entire Sample 321
Table J.3 Results of the Two-Level Cross-Classified Random Effects Model Predicting Students' Beliefs and Achievement Using the Single-Teacher Sample 322
Table J.4 Results of the Two-Level Cross-Classified Multiple Membership Random Intercept Model Predicting Students' Beliefs and Achievement Using the Entire Sample 323
Table L.1 Steps in the Coding Process and Prominent Changes 326
Table M.1 Coding Framework for Fixed-Ability Practices 329
Table M.2 Coding Framework for Motivational Beliefs 336
Table N.1 Example of Teacher Matrix Linking Teachers' Motivational Beliefs and other Justifications to Fixed-Ability Practices 339

Table N.2 Example of Student Matrix Linking Fixed-Ability Practices and Their Motivational Beliefs 343

List of Figures

Figure 2.1 Conceptual Framework for the Present Study 5
Figure 3.1 Graphical Representation of the Study Design 37
Figure 3.2 Flowchart of the Development of the Fixed-Ability Practices Questionnaires 53
Figure 4.1 Network Graph of the Cross-Classified Multiple Membership Structure 65
Figure 4.2 Steps in Examining the Dimensionality of Student-Perceived Fixed-Ability Practices 69
Figure 4.3 Simplified Graphical Representation of the Exploratory Structural Equation Models 72
Figure 7.1 Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Fixed Mindset 112
Figure 7.2 Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Mastery Goals 112
Figure 7.3 Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Performance-Approach Goals 112
Figure 7.4 Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Performance-Avoidance Goals 112
Figure 7.5 Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Failure-to-Ability Attribution 112
Figure 7.6 Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Self-Concept 112
Figure 7.7 Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Achievement 113
Figure 7.8 Marginal Effects Plot for Differential Value Predicting Failure-to-Ability Attribution 114
Figure 7.9 Density Plot for the Regression Coefficients of the Uniform Fixed-Ability Practices Dimension in the CC Model 115
Figure 7.10 Density Plot for the Regression Coefficients of the Uniform Fixed-Ability Practices Dimension in the CCMM Model 115
Figure 9.1 Thematic Map of Fixed-Ability Practices 127
Figure 9.2 Thematic Map of Motivational Beliefs 132
Figure J.1 Kernel Density Plot for Fixed Mindset in the CC Model 315
Figure J.2 Boxplot for Fixed Mindset in the CC Model 315
Figure J.3 Kernel Density Plot for Mastery Goal in the CC Model 315
Figure J.4 Boxplot for Mastery Goal in the CC Model 315
Figure J.5 Kernel Density Plot for Performance-Approach Goal in the CC Model 316
Figure J.6 Boxplot for Performance-Approach Goal in the CC Model 316
Figure J.7 Kernel Density Plot for Performance-Avoidance Goal in the CC Model 316
Figure J.8 Boxplot for Performance-Avoidance Goal in the CC Model 316
Figure J.9 Kernel Density Plot for Self-Concept in the CC Model 316
Figure J.10 Boxplot for Self-Concept in the CC Model 316
Figure J.11 Kernel Density Plot for Achievement in the CC Model 317
Figure J.12 Boxplot for Achievement in the CC Model 317
Figure J.13 Histogram for Failure-to-Ability Attribution in the CC Model 317
Figure J.14 Kernel Density Plot for Fixed Mindset in the CCMM Model 318
Figure J.15 Boxplot for Fixed Mindset in the CCMM Model 318
Figure J.16 Kernel Density Plot for Mastery Goal in the CCMM Model 318
Figure J.17 Boxplot for Mastery Goal in the CCMM Model 318

Figure J.18 Kernel Density Plot for Performance-Approach Goal in the CCMM Model 319
Figure J.19 Boxplot for Performance-Approach Goal in the CCMM Model 319
Figure J.20 Kernel Density Plot for Performance-Avoidance Goal in the CCMM Model 319
Figure J.21 Boxplot for Performance-Avoidance Goal in the CCMM Model 319
Figure J.22 Kernel Density Plot for Self-Concept in the CCMM Model 319
Figure J.23 Boxplot for Self-Concept in the CCMM Model 319
Figure J.24 Kernel Density Plot for Achievement in the CCMM Model 320
Figure J.25 Boxplot for Achievement in the CCMM Model 320
Figure J.26 Histogram for Failure-to-Ability Attribution in the CCMM Model 320

Chapter 1
Introduction

1.1 Background of this thesis

In the late 1980s, Dweck and Leggett (1988) developed a theoretical framework of individuals' "theories of intelligence, that is, their implicit conceptions about the nature of ability" (p. 262). Two implicit theories of ability were distinguished: an incremental theory and an entity theory, also referred to as a growth and fixed mindset, respectively (Dweck, 1999). Individuals with a growth mindset believe that ability can be expanded, whereas those with a fixed mindset believe that although new knowledge can be acquired, ability cannot be cultivated. These beliefs are thought to fall on a continuum from a strong fixed to a strong growth mindset (Dweck, 1999). Grounded in research on social cognition, the framework illustrates how mindsets are related to individuals' affect, cognition, and even behaviour (Dweck, 1999). Mindsets can give rise to a meaning system consisting of intertwined achievement-related motivational beliefs including achievement goal orientation (Payne et al., 2007), achievement attribution (Blackwell et al., 2007), and ability self-concept (Gunderson et al., 2017). Ultimately, mindsets can even impact achievement (Blackwell et al., 2007). Decades of research have demonstrated that holding a growth mindset is largely associated with adaptive outcomes, whereas holding a fixed mindset tends to be maladaptive (Burnette et al., 2013; Dweck & Leggett, 1988; Dweck & Molden, 2017). Given the centrality of mindset in this network of constructs, understanding how it develops is a great concern for educational psychology (see Haimovitz & Dweck, 2017) and the overarching purpose of the present study.

The importance of students' mindsets has been largely established, but relatively little is known about how they develop. It has been suggested that students might adopt a fixed mindset when they experience an environment that views ability as stable (Leondari & Gialamas, 2002). This is likely the case when students transition to secondary schools, which tend to be characterised by a greater emphasis on ability and performance rather than effort, fewer opportunities for student decision-making, and less emotional support (for a review, see Eccles & Roeser, 2009). In line with this, secondary school students tend to endorse fixed mindsets more strongly than primary school students (Cheng & Hau, 2003; Leondari & Gialamas, 2002), and motivation more generally declines during this transitional period (see Eccles & Roeser, 2009).

Students’ beliefs about ability and achievement may be influenced by their experiences in school because it is a context in which achievement is put in the spotlight. In particular, teachers guide students’ learning and are thus likely an essential source for their achievement-related motivational beliefs. While it is not necessarily assumed that teachers pass on their own mindset through direct socialisation processes, their specific behaviours may affect students’ beliefs (for a review, see Haimovitz & Dweck, 2017; Park et al., 2016). Hence, examining the role of teaching practices in the development of student mindset seems like a promising line of research. While multiple mindset teaching frameworks have been proposed (Boaler, 2013; Rissanen et al., 2019; Sun, 2018), only a few studies have directly investigated the impact of practices on mindset (e.g., Ommundsen, 2001; Park et al., 2016). More research is thus needed to examine this relationship.

Besides the need to understand what teaching practices foster students’ fixed mindsets, the question arises why some teachers engage in practices that are likely to encourage such fixed views of ability. Much research indicates that teaching practices are influenced by teachers’ beliefs, although these relationships are not always straightforward (e.g., for reviews, see Fang, 1996; Kagan, 1992; Pajares, 1992). Teacher beliefs can be broadly defined as “tacit, often unconsciously held assumptions about students, classrooms, and the academic material taught” (Kagan, 1992, p. 65). Due to the complexity of the school environment, teachers’ beliefs and behaviours may not always align; time and resource constraints as well as school policies and the need to prepare students for exams may affect their practices considerably (Fang, 1996). Nevertheless, few would contend that teachers’ beliefs affect their perceptions and judgments of students, which, in turn, influence their behaviour in the classroom (Fang, 1996). It has even been proposed that beliefs will be the most valuable psychological construct in teacher education (Pintrich, 1990) and that changing teacher beliefs is imperative for changing practices (Fives & Buehl, 2012). It is thus a reasonable proposition that teachers’ mindsets and related motivational beliefs may influence the teaching practices they engage in. Some studies indeed found that teachers with a fixed compared to growth mindset are more likely to engage in practices that can foster students’ fixed mindsets (e.g., Lee, 1996; Rattan et al., 2012; Stipek et al., 2001).

Fixed mindsets amongst students and teachers are particularly prevalent in mathematics (Gunderson et al., 2017; Jonsson et al., 2012). People may readily accept the myth that an individual is just not a “maths person,” which can create a barrier to maths success (see Chestnut et al., 2018). The belief that someone is either good at maths or not appears especially widespread in Western countries, where a greater emphasis is placed on ability when explaining maths performance compared to Asian countries (Stevenson et al., 1990). Interestingly, a study showed that maths-specific mindsets influenced motivation and achievement in maths, while writing- and reading-specific mindsets did not influence motivation and achievement in writing and reading (Gunderson et al., 2017). The authors thus argued that maths-specific beliefs are particularly important targets for intervention. Maths generally receives much attention in motivation research, potentially because students show relatively low interest and engagement in this subject. For instance, only half of Year 5 students (aged 9–10 years) and 14% of Year 9 students (aged 13–14 years) in England indicate that they very much like to learn maths (Greany et al., 2016). Moreover, although maths performance of students in the United Kingdom (UK) has improved over recent years, the UK significantly falls behind East Asian countries (OECD, 2019). Fixed mindsets may contribute to the problem of relatively low engagement and achievement in maths (Bostwick et al., 2019; Bostwick et al., 2020), which shows the need for fostering adaptive motivational beliefs in this subject.

1.2 The present study

Considering existing gaps in the literature, the overarching purpose of the present study was to understand the role of teaching practices in the development of students' mindsets and associated achievement-related motivational beliefs, namely achievement goal orientation, achievement attribution, and ability self-concept. I first aimed to develop a conceptualisation of, and a teacher and student questionnaire for, the assessment of fixed-ability practices (FAPs) – teaching practices that are based on the notion that ability is fixed. I then aimed to examine what the different dimensions of student-perceived FAPs are (research question (RQ) 1). Note that only student reports were used to determine the factor structure of FAPs as it was not feasible to recruit a sufficiently large teacher sample. Next, I aimed to examine whether teachers' motivational beliefs predict their engagement in student-perceived and/or teacher-reported FAPs (RQ 2). Moreover, I aimed to examine whether student-perceived and/or teacher-reported FAPs predict students' motivational beliefs and achievement (RQ 3). To gain a more qualitative understanding of these relationships, I explored how teachers justify their engagement in FAPs, and whether these justifications are related to their motivational beliefs (RQ 4), as well as how students perceive FAPs, and what messages they draw from FAPs for their motivational beliefs (RQ 5).

To answer these research questions, a longitudinal mixed-methods design was used. Teachers' and students' motivational beliefs in maths were quantitatively assessed with questionnaires at the beginning of the first year of secondary school (Year 7), and students' motivational beliefs were again assessed at the end of the school year. During the middle of the year, FAPs were quantitatively assessed with teacher and student questionnaires. I also conducted classroom observations during maths lessons of four selected teachers and used video-stimulated recall interviews with these teachers and some of their students to encourage reflections on FAPs. Students' maths grades were obtained from school records.

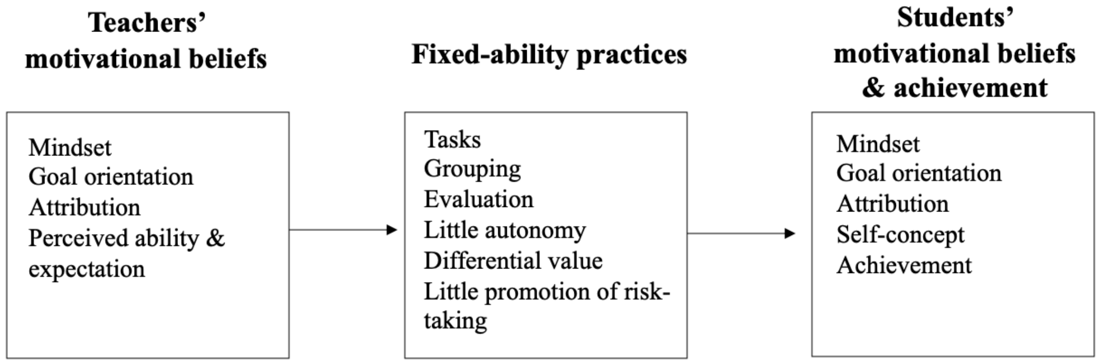
This thesis is divided into twelve chapters including this introduction. Chapter 2 reviews the literature and thus situates the study in the field of educational psychology, and, in particular motivation. The review identifies gaps in the literature and lays the foundations for the research questions and the conceptual framework illustrating the hypothesised relationships. Chapter 3 describes and justifies the methodology of the study. It provides a rationale for the research design and methods, introduces the teacher and student sample, describes the selection and development of specific research instruments, and presents the procedures of the main study. Chapter 4 describes the general quantitative analytical approach as well as the analysis of the dimensionality of FAPs (RQ 1). Results for RQ 1 are then presented in Chapter 5. This is followed by Chapter 6, which describes how the relationships between teachers' motivational beliefs and their engagement in FAPs (RQ 2) and the relationships between FAPs and students' motivational beliefs and achievement (RQ 3) were analysed. The results of these analyses are subsequently presented in Chapter 7. Turning to the qualitative data, Chapter 8 describes how the teacher and student interviews were analysed. Chapter 9 presents the themes that were identified in interviews as well as the findings relating to how teachers justify their engagement in FAPs (RQ 4). Chapter 10 then presents the findings relating to what messages students draw from FAPs for their motivational beliefs (RQ 5). This is followed by Chapter 11, which discusses the results. More specifically, it summarises and evaluates the main findings and situates them in the current body of literature. Finally, Chapter 12 concludes this thesis. It highlights this study's contributions and implications for educational practice and research, acknowledges its limitations, and proposes avenues for future research.

Chapter 2
Literature Review

This literature review has three sections. The first section (Section 2.1) reviews research on the importance of mindset for learning and positions mindset at the centre of a meaning system of intertwined achievement-related motivational beliefs. In the second section (Section 2.2), the largest section of this chapter, I start to conceptualise fixed-ability practices (FAPs). I review research on teachers' and students' mindsets in relation to teaching practices, and sometimes also draw on studies about parents' and children's mindsets in relation to parenting practices because these might share common mechanisms. I also review research suggesting that FAPs might adversely affect students' achievement. In the last section (Section 2.3), I highlight the need for the present study and present the research questions. Research articles included in this review were selected based on literature searches with ERIC, PsycINFO, and Google Scholar. The selection of key terms[1] for the initial search was informed by my previous knowledge of studies on mindset and motivational beliefs. I later manually sought papers cited in the reference sections of relevant articles and added new key terms to learn more about specific beliefs or teaching practices. To provide an overview of all constructs of interest, the conceptual framework for the present study is presented in Figure 2.1.

Figure 2.1

Conceptual Framework for the Present Study



[1] Initial search terms included: ability differentiation, achievement goal orientation, achievement attribution, classroom goal structure, differential treatment, expectancy effect, implicit theories of ability/intelligence, instructional practices, mindset, self-concept, teacher beliefs, teacher expectation, teaching methods, teacher mindset, and teaching practices.

2.1 The meaning system of mindset and related motivational beliefs

Whether or not ability can be changed is an important motivational belief that can give rise to a meaning system of other achievement-related motivational beliefs and even affect achievement as will be shown in this section. Holding a growth mindset rather than a fixed mindset is positively associated with a range of standardised performance measures as well as school grades in reading and maths across different age groups and countries (e.g., Blackwell et al., 2007; Cury et al., 2006; De Castella & Byrne, 2015; Diseth et al., 2014; McCutchen et al., 2016; Lüftenegger et al., 2015). While some studies did not find such associations (e.g., Bahník & Vranka, 2017; Leondari & Gialamas, 2002), meta-analyses found a significant, if only weak, direct positive association between growth mindset and academic achievement ($r = .095$, Burnette et al., 2013; $r = .07$, Costa & Faria, 2018; $r = .10$, Sisk et al., 2018). Potential explanations for relatively low effect sizes are that study results are heterogenous, and that associations are generally stronger among people facing academic difficulties (Yeager & Dweck, 2020). Importantly, some studies used a mindset intervention and found that students who developed a stronger growth mindset also showed improved achievement, which suggests that the relationship between mindset and achievement might be causal in nature (e.g., Aronson et al., 2002; Blackwell et al., 2007; Good et al., 2003).

Besides impacting achievement, mindsets can give rise to a range of other important motivational beliefs (Dweck, 1999). The present study focused on achievement goal orientations, achievement attributions, students' ability self-concept, and teachers' perceived ability/expectation for their students. Each of these constructs and their relationship with mindset will be discussed below. Most of what we know about the relationships between mindset and motivational beliefs is based on research with students. However, more recently, there has been an increased interest in investigating mindsets amongst teachers (e.g., Butler, 2000; Patrick et al., 2001; Rattan et al., 2012). Goal orientations describe students' motivation to perform well in school and teachers' motivation to be a good teacher. Achievement attribution, students' ability self-concept, and teachers' perceived student ability/expectation describe how students and teachers explain and perceive students' ability and achievement. Teachers' goal orientations somewhat stand out from the other selected teacher constructs because they do not relate to what teachers believe about their students. Yet, they were included in this study because they have been identified as an important influence on the teaching practices of interest. All these motivational beliefs are interrelated but conceptually different. Hence, a more accurate understanding of the relationships between teaching practices and motivational beliefs is enabled when considering these beliefs simultaneously.

It is important to highlight that these motivational beliefs are domain-specific and are thus treated as such throughout the study. Students tend to have a stronger fixed mindset in maths compared to reading and writing (Gunderson et al., 2017). Similarly, language and social science teachers have stronger growth than fixed mindsets, whereas maths and science teachers show no such preference (Jonsson et al., 2012). Regarding achievement goal orientations, mastery goals have only low to moderate correlations between different subjects, but performance goals are strongly correlated between subjects (Bong, 2004). Achievement attributions have also been found to differ greatly between subjects (Bong, 2004). Lastly, ability self-concept is domain-specific, and important domain comparison effects exist, so that, for example, higher performance in maths is typically negatively related to ability self-concept in more verbal subjects (Möller et al., 2020). In the following, I review the evidence linking mindset to students' and teachers' goal orientations and achievement attributions, students' ability self-concept, and teachers' perceived student ability/expectation. This will demonstrate the centrality of mindset in a network of motivational beliefs that are important for students' learning.

2.1.1 Achievement goal orientation

Three achievement goals for students are typically differentiated: A mastery goal (increasing one's understanding), a performance-approach goal (demonstrating one's ability), and a performance-avoidance goal (avoiding appearing incompetent; for reviews, see Elliot, 2005; Senko et al., 2008). Mastery-approach goals have also been distinguished from mastery-avoidance goals, the striving to avoid incompetence in intrapersonal rather than normative terms (Elliot &McGregor, 2001). Compared to items assessing the other three goals, mastery-avoidance items contain no goal-relevant language but refer to concerns and fears. Following Hulleman et al. (2010), I think it is helpful to use a more constrained definition of achievement goals and thus focus only on the former three concepts, particularly because mastery-avoidance goals are rarely spontaneously reported (Lee & Bong, 2016). I thus used the trichotomous goal orientation framework. Mastery goals are linked to a range of adaptive behaviours in school such as using better learning strategies, whereas performance-avoidance goals are linked to undesirable behaviours such as not seeking feedback as well as lower performance (for reviews, see Harackiewicz et al., 2002; Hulleman et al., 2010; Payne et al., 2007; Senko et al., 2008). The consequences of endorsing performance-approach goals are mixed, which can be explained by different operationalisations: some studies use items focusing on outperforming others (i.e., normative goals) and some studies use items focusing on appearing competent (i.e., appearance goals). The latter is generally related to detrimental behaviours, including self-handicapping and avoiding help, as well as lower performance (Hulleman et al., 2010; Senko & Dawson, 2017). I thus focused on such appearance goals in the present study. A meta-analysis shows that holding a fixed mindset is negatively associated with having mastery goals ($r = -.12$, but positively with having performance-avoidance goals ($r = .09$) and performance-approach goals ($r = .10$, Payne et al., 2007).

Teachers also hold achievement goals since schools are an achievement context for them as well (Butler, 2007). Similar to the conceptualisation of students' goal orientations, teachers may aim to develop their teaching competence, to demonstrate high competence or to avoid demonstrating inferior competence (Shim et al., 2013). Teachers with a fixed mindset tend to have weaker mastery goals for their teaching (Shim et al., 2013). While Shim et al. found no relationship between fixed mindset and performance goals, another study found a positive relationship between fixed mindset and aiming to be better than others (normative-approach) and not be worse than others (normative-avoidance; Mascret et al., 2017).

2.1.2 Achievement attribution

Successes and failures can be attributed to either internal factors, such as ability or effort, or external factors, such as situation or luck (Weiner, 1985). Such achievement attributions are important because they have behavioural consequences. Particularly maladaptive are failure-to-ability attributions as students who believe that their poor performance is mostly explained by their lack of ability rather than effort tend to respond in helpless ways and stop trying (e.g., Robins & Pals, 2002). Previous research showed that students with fixed mindsets tend to make more ability than effort attributions, whereas students with growth mindsets tend to attribute performance more to effort than ability (e.g., Blackwell et al., 2007; Hong et al., 1999; Robins & Pals, 2002).

Teachers also tend to generate explanations for their students' performance. Student ability and effort belong to the most common attributions that teachers make for their students' achievement (Fennema et al., 1990; Georgiou et al., 2002; Natale et al., 2009). Rattan et al. (2012) showed that these attributions might be related to teachers' mindsets in a study with undergraduate students who were asked to imagine they were 7th-grade maths teachers and graduate students who taught undergraduates in maths-related fields. Participants with fixed mindsets were more likely than those with growth mindsets to attribute a hypothetical student's performance to "not being smart enough" or a "lack of maths intelligence" rather than to a “lack of effort.” This also occurred when mindsets were experimentally induced. In another study, Butler (2000) presented teachers and students with a scenario in which a student either improved or became worse in maths over time. She found that those with fixed mindsets attributed the earlier performance more than the later one to ability, and the earlier performance less than the later one to luck. Such temporal effects did not occur amongst individuals with growth mindsets. Moreover, those with fixed mindsets tended to judge students as having higher ability in the deteriorating compared to improving scenario, whereas the opposite was the case for those with growth mindsets.

2.1.3 Students’ ability self-concept and teachers’ perceived student ability/expectation

First, it is important to define ability self-concept and explain why I chose it over similar constructs, such as academic self-concept and self-efficacy. The term ability self-concept, sometimes also termed ability self-perception, comes from Eccles et al. (2005). Broadly speaking, maths ability self-concept answers the question “How good in maths are you?” (Eccles et al., 2005) and can thus be described as a subjective, cognitive appraisal of a person’s current ability. Theoretically, this can be distinguished from a person’s future performance expectation; however, these two constructs are empirically indistinguishable and load on a single factor (Eccles & Wigfield, 1995). Thus, evaluations of current performance and future expectations are often treated interchangeably (e.g., Jussim & Eccles, 1992; Bohlmann & Weinstein, 2013; Marsh et al., 2018), which was also done in the present study.

Ability self-concept can be distinguished from academic self-concept, which refers to individuals’ general perceptions of themselves in an academic subject. Besides perceived ability or competence, it commonly includes the components “interest in” and “perceived value of” a given domain (for a review, see Bong & Skaalvik, 2003). While academic self-concept is thus more inclusive, the current study was interested in achievement-related motivational beliefs, which is why the cognitive appraisal of ability was more relevant than affective evaluations pertaining to maths. Ability self-concept also needs to be distinguished from self-efficacy because both are conceptualised as a person’s perception of their competence. Bandura (1997) defines self-efficacy as one’s subjectively perceived competence to produce given achievements. This is usually assessed by asking participants how confident they are to solve various tasks of different difficulty related to the same subject (Bandura, 2006). It thus usually measures confidence to complete specific tasks rather than an overall judgement of how good students think they are (or teachers think students are) at maths. Note that some self-efficacy measures deviate from this description, which Marsh et al. (2018)

would not classify as pure self-efficacy measures but as more self-concept-like. There are additional relevant differences. Ability self-concept or perceived ability are usually measured by asking individuals to rate themselves or others compared to other students in the class, respectively (Eccles et al., 1993; Eccles et al., 2005), whereas self-efficacy items do not rely on social comparison (Marsh et al., 2018). It was critical to take such a reference group into account in the present study because students tend to learn about their ability through teachers' differential treatment of low- and high-performing students as well as public performance comparisons (Bohlmann & Weinstein, 2013; Dickhäuser et al., 2017). Please note that after having defined ability self-concept, I will simply refer to "self-concept" throughout the remainder of this thesis.

Self-concept has positive reciprocal relations with achievement and is as such an important driver of learning processes (Sewasew & Schroeders, 2019; Shanley et al., 2019). Research showed that students' growth mindsets are positively related to perceived performance (Ahmavaara & Houston, 2007; Gonida et al., 2006) and that students with fixed mindsets tend to feel less competent or efficacious in maths (Gunderson et al., 2017; Zander et al., 2018). Yet, the direction of influence is not clear as mindset might also be the result of perceived competence (Gonida et al., 2006). Moreover, other studies found no relationship between students' mindset and perceived competence (Cury et al., 2006; Leondari & Gialamas, 2002). From a theoretical perspective, students with fixed mindsets tend to infer their ability from their performance and should thus perceive lower ability as well as expect lower future performance following failure (Dweck, 1999).

Research suggests that teachers' mindsets are related to their expectations for students' performance. Butler (2000) demonstrated that teachers with fixed rather than growth mindsets tend to base their ability judgements on students' initial performance without adjusting appraisals when later performance diverges. In contrast, those with growth mindsets might be more likely to adjust their ability judgement according to performance changes. Butler thus suggested that there would be greater expectancy effects for teachers with fixed mindsets. Expectancy effects occur when teachers' expectations affect students' actual performance due to direct effects, i.e., because teachers adjust the curriculum according to expectations, and indirect effects, i.e., because teachers influence students' perceived ability, expectations, and behaviours, which can affect their performance (Bohlmann & Weinstein, 2013; Good & Nichols, 2001). A relationship between teacher mindset and expectations was supported by a study that found that graduate students instructing undergraduates had lower expectations for hypothetical failing students when they had fixed mindsets compared to growth mindsets

(Rattan et al., 2012).

This section reviewed studies that demonstrated that students' and teachers' mindsets are central to a range of other important motivational beliefs, as well as that students' mindsets can affect their achievement. Because of the close links between these different motivational beliefs, it was important to consider these constructs in tandem when investigating the effects of fixed-ability practices.

2.2 A conceptualisation of fixed-ability practices

This section reviews literature on teaching practices that are based on the conception that ability is fixed. I focused on fixed-ability practices rather than adaptive practices because students appear to start their schooling with adaptive beliefs but then tend to adopt more fixed mindsets when transitioning to secondary school, presumably because of the exposure to maladaptive practices as discussed in the introduction. Accordingly, identifying such maladaptive practices and, ultimately, reducing teachers' use of these practices may prevent the development of such detrimental beliefs.

To identify FAPs, I mainly synthesised the literature from two areas: a) relationships between teachers' motivational beliefs and their practices, and b) relationships between teaching practices and students' motivational beliefs. In addition, I sometimes drew on research about c) relationships between parents' motivational beliefs and their parenting practices, and d) relationships between parenting practices and children's motivational beliefs. While it is not certain that the mechanism relating to parenting practices directly transfer to teaching practices, parenting practices can at least point towards potential important teaching practices when such practices can be used by both parents and teachers. It thus appeared important to consider these, especially if relatively little prior research on relationships between a given teaching practice and students' beliefs was available. Please note that the literature review mostly focused on mindset although the other motivational beliefs were also examined.

Throughout this thesis, teaching practices refer to all teacher behaviours and verbal statements during lessons. Although all teaching practices can be viewed from a mindset perspective (Boaler, 2013), I selected practices that have previously been empirically and/or theoretically linked to students' and/or teachers' mindsets. After identifying such practices, I grouped conceptually related practices into dimensions. This finally resulted in six broad practice dimensions with subdimensions, namely 1) tasks (unsophisticated task differentiation, unidimensional tasks), 2) grouping (ability grouping, stability), 3) evaluation (performance over process recognition, social comparison), 4) little autonomy (little choice, little promotion of self-regulation), 5) differential value, and 6) little promotion of risk-taking (little encouragement of help-seeking, unconstructive handling of mistakes).

An overview of these practices is provided in Table 2.1. Each dimension comprises practices that can be characterised as either fixed-ability practices, which were assumed to foster fixed mindsets and related maladaptive motivational beliefs, or non-fixed-ability practices, which were expected to foster growth mindsets and other adaptive motivational beliefs or at the least not to foster maladaptive motivational beliefs. The table pitches FAPs and non-FAPs against each other, whereby the indicators of subdimensions often represent opposites of each other (e.g., "different difficulty/amount for some" vs. "same difficulty/amount for all" and "visible/public evaluation" vs. "private evaluation"). It is important to note that these practices should be regarded as the extreme ends of a *continuum* rather than as a dichotomy of "either/or". While individual practices can theoretically be mapped onto these two broader teaching approaches, how these are being actualised in the classroom is likely much more complex than that, and it may not always be clear where the boundaries between FAPs and non-FAPs lay. Teachers may differ in the frequency with which they engage in these practices, and it seems unlikely that a teacher would always engage in only FAPs or non-FAPs across all dimensions. Yet, based on theoretical considerations and empirical work reviewed below, teachers with a fixed mindset may generally lean towards fixed-ability practices, or at least to a greater extent than those with a growth mindset.

The present dimensions are distinct from, but also overlap with, other mindset teaching frameworks. Boaler (2013) also proposed six practice dimensions that can convey mindset messages: questions asked, tasks, grouping, norm-setting, grading and feedback, and treatment of mistakes, but she only discussed grouping practices and treatment of mistakes in more detail. Two other mindset teaching frameworks, one by Rissanen et al. (2019) and one by Sun (2018), have been published after the data collection of the present study. These also show some overlap, but the present study makes important additions, which is elaborated on in the Discussion (Chapter 11). Importantly, no quantitative measurement tool exists for these frameworks.

Table 2.1

Dimensions and Characteristics of Fixed-Ability Practices

Dimension	Sub-dimension	Fixed-ability practices	Non-fixed-ability practices
Tasks	Unsophisticated task differentiation	Different tasks for some	Same tasks for all
		Different difficulty/amount for some	Same difficulty/amount for all
		Challenge for some	Challenge for all
	Unidimensional tasks	Only one approach/solution/product	Different approaches/solutions/products
		No extension tasks	Extension tasks
Grouping	Ability grouping	Homogenous ability groups/pairs	Heterogeneous ability groups/pairs
		Groups teacher dictated	Group choice
	Stability	Fixed groups	Fluid groups
Evaluation	Performance over process recognition	Performance recognised without effort reference	Process and effort recognised
		Only one try for assignments	Chance to improve work
	Social comparison	Competitive rewards	Non-competitive rewards
		Visible/public evaluation	Private evaluation
		Competition encouraged	Little competition
Little autonomy	Little choice	No choice over task selection and approach	Choice over task selection and approach
		No choice over work partner	Choice over work partner
	Little promotion of self-regulation	No student self-evaluation	Student self-evaluation
		No student self- and/or peer-assessment	Student self- and/or peer-assessment
Differential value		Critical/dismissive towards some	Accepting/warm towards all
		Put-downs are tolerated	Compliments encouraged
		Listen more to some	Listen carefully to all
		Some participate	All participate
		Threatening (e.g., extra homework)	Non-threatening
Little promotion of risk-taking	Little encouragement of help-seeking	Peer and teacher help-seeking discouraged	Peer and teacher help-seeking encouraged
	Unconstructive handling of mistakes	Mistakes as debilitating	Mistakes as enhancing
		Critical of mistakes or ignore mistakes	Discuss and use mistakes for learning

Perhaps not surprisingly, the present dimensions also resemble the TARGET framework (Ames, 1992a, 1992b; Epstein, 1988), which includes six instructional dimensions that can foster a mastery goal in the classroom (task, authority, recognition, grouping, evaluation, and time). Although this framework is well established and corresponding measurements exist (e.g., Lüftenegger et al., 2017; Vedder-Weiss & Fortus, 2017), mindsets and goal orientations are not identical concepts. Thus, a fresh perspective and synthesis of works related to mindset were warranted. This becomes especially apparent when looking beyond the naming of dimensions because several dimensions may seem similar but are conceptualised quite differently. For instance, the TARGET grouping dimension focuses on whether students work in groups while the FAPs grouping dimension focuses on the bases for and stability of grouping. Moreover, the TARGET framework does not include a dimension on differential value. These and other differences will be expanded on in the Discussion (see Section 11.1).

Another framework that these dimensions overlap with is the ability-differentiated practice framework developed by Bohlmann and Weinstein (2013). Their framework comprises practices that can convey teacher expectations to students, which includes the dimensions curriculum, grouping, motivational strategies, evaluation, teacher control/choice, and climate. Teacher expectations are also of interest in the present study and are related to mindset as reviewed above (see Section 2.1.3), which may explain the similarities between the conceptualisations. Importantly, no questionnaire exists for their framework because Bohlmann and Weinstein (2013) assessed these dimensions in a qualitative study.

In the following, I describe each FAPs dimension and review studies that revealed or suggested links between the practice and students' and/or teachers' mindset as well as related motivational beliefs. Please note that this review also includes studies that have been published after I developed the FAPs conceptualisation and corresponding questionnaire, which are thus not included in Table 2.1 (e.g., Barger, 2019; Rissanen et al., 2019; Sun, 2019). Some of this work will show that the relationships between teachers' mindset and FAPs as well as FAPs and students' mindset are complex and not necessarily always in line with theory and my expectations. This will be explored more deeply in the discussion and conclusion.

2.2.1 Tasks

The task dimension encompasses two aspects; first, *unsophisticated task differentiation* (e.g., giving fewer and easier tasks to low-performing students) and, second, setting *unidimensional* *tasks* (e.g., tasks that can be solved in only one way and have only one correct solution). When teachers assign tasks depending on students' ability, they provide an ability-differentiated curriculum (Bohlmann & Weinstein, 2013). And when they assign tasks that can be solved in only one way, they use unidimensional tasks (Boaler, 2016). In contrast, multidimensional tasks can be solved with different methods and are accessible to all students (Boaler, 2016). These tasks become increasingly more challenging, so that they have a low threshold but high ceiling. Boaler (2016) argues that such tasks are part of a growth mindset pedagogy, while unidimensional tasks allow only some students to be successful.

2.2.1.1 Teachers' motivational beliefs and tasks

Multiple studies suggest that teachers with fixed mindsets provide students with less access to the curriculum through unsophisticated task differentiation. It has been demonstrated that fixed mindset teachers are more likely than growth mindset teachers to assign students easier tasks following poor maths performance, assuming that experiencing success would motivate students and make them feel better (Lee, 1996; Rissanen et al., 2018). In contrast, teachers with growth mindsets tended to assign more difficult tasks to help students improve their skills. Similarly, it has been found that fixed mindset undergraduates, who imagined to be maths teachers, were more likely to report that they would assign poorly performing students with less homework (Rattan et al., 2012). Assigning multidimensional tasks is more in line with a growth mindset (Boaler, 2016). Yet, Sun (2019) revealed through observations and interviews with growth mindset teachers that not all used such tasks consistently and instead insisted on following step-by-step procedures. Sun (2019) also found that some growth mindset teachers communicated the expectation that not all students were able to solve certain tasks or use certain methods. This indicates that the relationships between teachers' mindset and their practices might not be straightforward. However, drawing strong conclusions from this study is difficult because it included no comparison group of fixed mindset teachers.

Previous research showed that teachers with low expectations are more likely to report that they assign discrete activities for high- and low-ability students than teachers with high expectations, and that they favour more curriculum differentiation (Rubie, 2003). Because of low expectations, teachers tend to provide students in lower-ability groups with less access to the curriculum, whereas they provide similar curriculum access for students in mixed-ability groups (Ireson & Hallam, 2001 as cited by Ireson et al., 2002; Harris, 2012; Standholtz et al., 2004). Teachers differentiate the curriculum amongst other reasons because they think that low intelligence prevents students from meeting expected standards (Harris, 2012). This points towards a relationship between attributing performance to ability and curriculum differentiation. However, no studies were identified that examined this directly. Neither were studies found that investigated associations between teachers' goal orientations and their task assignment. From a theoretical perspective, teachers with performance goals would be more likely to differentiate the curriculum because this would promote the best students and prevent weaker students from experiencing failure. This conforms to the correspondence hypothesis, which holds that teachers apply goals to students that they hold for themselves (Dresel et al., 2013). It would also agree with the functionality hypothesis, stating that teachers use practices that are most likely to fulfil their own goals (Dresel et al., 2013), because teachers with performance goals may appear more competent if their students can solve the tasks they are given compared to when they struggle. In addition, assigning tasks with one correct solution that need to be solved in one specific way would make it easier for teachers to compare students' performance.

2.2.1.2 Tasks and students' motivational beliefs and achievement

I found no studies that examined the relationship between task assignment and students' mindsets. However, it seems plausible that if some students get easier or less work while others get more difficult tasks, students may come to believe that only some people have the ability to solve such difficult tasks. Similarly, Boaler (2016) argued that unidimensional tasks can send the message that only a few "able" students can access maths and that maths is only about finding the one correct answer to a problem. In line with this, it has been argued that students "come to see mathematics as a fixed and closed subject" when they predominantly engage in unidimensional tasks, whereas they come "to see mathematics as an open and growth subject" when they engage in tasks with multiple entry points that allow for different methods (Anderson et al., 2018, p. 3). I found no studies examining task assignment and students' goal orientation and attribution, but whether a teacher assigns students with easier or more difficult tasks may convey teachers' expectations, which, in turn, is likely to influence students' self-concept. Undergraduate students reported that they would have lower expectations if their teachers were to assign them less and easier work (Rattan et al., 2012). Similarly, students who are underestimated by their teachers tend to have lower perceived competence and expectations compared to overestimated students (Zhou & Urhahne, 2013). This might be explained by FAPs because Kuklinski and Weinstein (2001) showed in a longitudinal study that teachers' expectations predicted students' expectations among 5th-graders, but only in highly ability-differentiated classrooms. Likewise, in classrooms with more observed ability-based differentiation, students tended to rate their maths ability lower compared to classrooms with little differentiation, even after controlling for teacher expectations (Bohlmann & Weinstein, 2013). Teachers' and students' expectations were also more congruent in highly differentiating classes.

I did not find direct evidence linking teachers' task assignment to students' achievement. Yet, effects of ability grouping on achievement may be due to task differentiation. For instance, higher maths achievement of students placed in high-ability groups compared to students with similar ability placed in middle or low-ability groups has been argued to be due to curriculum differences (Ireson et al., 2002). It has also been proposed that the influences of teacher expectations on students' achievement are due to task differentiation (Kuklinski & Weinstein, 2001). Moreover, it seems logical that receiving easier tasks will negatively influence achievement over time. Although not directly examined in this study, influences of tasks, and FAPs more generally, on achievement might also be mediated by the proposed changes in students' motivational beliefs as these can affect achievement as reviewed earlier (see Section 2.1).

2.2.1.3 Brief summary on tasks and mindset

Several studies indicated that teachers with fixed mindsets are more likely to engage in unsophisticated task differentiation. However, these studies used hypothetical designs (Lee, 1996; Rattan et al., 2012) or were only qualitative in nature (Rissanen et al., 2018). Moreover, a qualitative study showed that teachers with growth mindsets did not consistently use multidimensional tasks although these are theoretically in line with a growth mindset (Sun, 2019). The lack of a comparison group renders definite conclusions difficult. Influences of task assignment on students' mindsets have been supported theoretically (e.g., Anderson et al., 2018; Boaler, 2016) but empirical studies are lacking. Overall, it becomes very apparent that more research investigating links between task assignment and teachers' and students' motivational beliefs, in particular their mindset, is necessary.

2.2.2 Grouping

The grouping dimension reflects two aspects of in-class grouping practices: *ability grouping*, that is whether or not students are grouped or partnered with similarly performing students, and *stability*, that is whether grouping is flexible or fixed over time. There is a long tradition of grouping by ability in the UK, which includes the use of streaming (students with similar abilities stay in the same group for all subjects), banding (students are allocated to broader ability bands) and setting (students are placed in subject-specific ability groups). These practices are particularly prevalent in maths (Hallam et al., 2003). Although these forms of grouping practices are largely determined by school policies, individual teachers can usually decide whether or not to use within-class ability grouping (Blatchford et al., 2008). Accordingly, even in schools with a reputation for mixed-ability teaching, some teachers use ability grouping in their classrooms (Marks, 2013).

2.2.2.1 Teachers' motivational beliefs and grouping

Forming homogenous and stable ability groups has been argued to be based on the notion of fixed abilities (Boaler, 2013; Marks 2013). Similarly, it has been proposed that fixed mindset teachers may be more likely than growth mindset teachers to use ability grouping and to teach ability groups in different ways because they might be less willing to invest effort in children with a low perceived ability (Stipek et al., 2001). Indeed, research showed that teachers with fixed mindsets were more likely to recommend that a hypothetical student struggling with maths would be placed in a homogenous rather than heterogeneous ability group (Lee, 1996). The opposite was found for teachers with growth mindsets. A qualitative study also found that a fixed mindset teacher favoured homogenous ability grouping (Rissanen et al., 2018). In contrast to these findings, Sun (2019) demonstrated that it was common even for teachers with growth mindsets to form homogenous ability groups or to seat low- and high-performing students separately. This indicates that mindsets do not necessarily directly translate into practice.

It is known that teachers hold lower expectations for students in low-ability compared to high-ability groups (Boaler et al., 2000), but I found no studies that examined whether teachers with low expectations for some students and high expectations for others are more likely to arrange groups by ability. To the best of my knowledge, there are also no studies that linked teachers' goal orientations or achievement attributions to ability grouping. Since the idea underlying ability grouping is closely related to that of task differentiation, similar relationships as identified above may be expected. Hence, teachers who have performance goals or attribute performance to ability might use more ability grouping.

2.2.2.2 Grouping and students' motivational beliefs and achievement

Boaler (2013) and Marks (2013) argued that ability grouping communicates fixed mindset messages. Ability grouping gives the impression that only some students can be high achievers, especially if movement between groups is limited (Boaler et al., 2000). Moreover, Marks (2013) proposed that children in low-ability groups identify themselves "with the limited ability and mathematical identities available to that group, limiting how they could act and who they could be" (p. 5). Students' goal orientations may also be influenced by ability grouping as a study demonstrated that students set for maths scored significantly higher on performance-approach and -avoidance goals than unset students, presumably due to increased salience of social comparison and competition (Butler, 2008). Another study also showed that normative goals (i.e., aiming to outperform others) were particularly prevalent amongst students in ability sets (Lee & Bong, 2016). Butler (2008) found that setting was unrelated to mastery goals and noted that ability grouping aims at providing students with appropriate challenge, which should theoretically orient them towards mastery goals. Yet, she only found detrimental effects on goal orientation. Even though these two studies did not look at within-class grouping, they suggest that relationships between within-class grouping and motivational beliefs may exist. Furthermore, research has linked ability grouping to students' self-concept. For instance, students placed in low within-class ability-groups in maths lowered their expectations, whereas those placed in high-ability groups raised their expectations (Reuman, 1989). Additionally, students in low-ability groups seem to have little hope of moving upwards because they do not believe they currently receive the teaching necessary to perform well in higher groups (Boaler et al., 2000).

Findings in the UK generally suggest relatively little influence of school-wide ability grouping on students' achievement. However, when influences are found, ability grouping mainly benefits higher-achieving students and is detrimental to lower- and mid-range-performing students (Education Endowment Foundation, 2018; Higgins et al., 2013; Ireson et al., 2002).

2.2.2.3 Brief summary on grouping and mindset

Forming homogenous and stable ability groups has been theoretically linked to fixed-mindset beliefs (Boaler, 2013; Marks 2013; Stipek et al., 2001). Evidence for this comes from a study using hypothetical scenarios (Lee, 1996) and a purely qualitative study (Rissanen et al., 2018), which constitute a good starting point. Another qualitative study provided some contrasting findings as it found that even growth mindset teachers formed homogenous ability groups, but it lacked a control group (Sun, 2019). Ability grouping, especially stable grouping, has theoretically been argued to communicate fixed mindset messages to students (Boaler, 2013; Boaler et al., 2000; Marks, 2013). Overall, more empirical evidence is needed to support the theoretical links between ability grouping and teachers' and students' motivational beliefs.

2.2.3 Evaluation

This dimension refers to two aspects of evaluation practices; *performance over process recognition* (e.g., praising being clever rather than working hard), and *social comparison* (e.g., pointing out high performers and comparing students' performance). These practices are also at the heart of the distinction between mastery-oriented and performance-oriented classrooms (Ames, 1992a, 1992b). Hence, research on classroom goal structures (i.e., classroom climates characterised by a given goal orientation) and mastery- and performance-oriented practices are reviewed here. The measures commonly used to assess these, especially measures of performance-oriented practices, predominantly involve items relating to evaluation (see the Patterns of Adaptive Learning Strategies (PALS) survey, Midgley et al., 2000). Yet, reviewed studies may also touch upon other dimensions, which cannot be completely separated due to the nature of the questionnaires used. Mastery goal structures are characterised by an emphasis on learning, effort, and personal progress, whereas performance-oriented classrooms promote competition and social comparisons (Ames, 1992a, 1992b; Anderman & Wolters, 2006; Urdan & Turner, 2005). Even though personal performance goals are separated into approach and avoidance tendencies, these are generally not distinguished for goal structures (e.g., Deemer, 2004; Park et al., 2018; Shim et al., 2013; Wolters & Daugherty, 2007). While this is theoretically possible, performance-avoidance goal structures have failed to form a reliable scale in the past, and performance-approach goal structures have been found to foster both performance-approach and -avoidance goals (Wolters, 2004). Such a distinction might thus not add explanatory value, which is why I followed previous studies by contrasting only mastery and performance goal structures. Notably, teachers do not necessarily promote only one of these goals (Patrick et al., 2001).

2.2.3.1 Teachers' motivational beliefs and evaluation

Teachers' fixed mindsets tend to positively relate to their self-reported performance-oriented practices, whereas their growth mindsets tend to positively relate to mastery-oriented practices (Miele et al., 2019; Park et al., 2016). Teacher growth mindset was also positively associated with student-reported mastery-oriented goal structures (Dresel et al., 2013). In addition, observations showed that emphasising performance was more characteristic of maths teachers with fixed compared to growth mindsets (Stipek et al., 2001). Similarly, university lecturers with fixed mindsets tend to be less likely than those with growth mindsets to emphasise learning and development as reported by students (Canning et al., 2019). Yet, Shim et al. (2013) found no significant relationship between teachers' mindset and self-reported goal structure although fixed mindset was negatively related to their mastery goals for teaching and goals for teaching were strongly related to goal structures.

More evidence for a relationship between mindset and evaluation practices comes from research with parents. A study showed that parents who believed their children's general and/or maths ability was fixed were more likely to report that they engaged in performance-oriented behaviours (e.g., expressing frustration towards the child) and less likely to engage in mastery-oriented behaviours (e.g., praising the child for working hard) when they helped their child with challenging maths tasks compared to parents who believed their children's ability could be changed (the same was found for verbal ability and verbal tasks; Muenks et al., 2015). However, another study found that parents with growth mindsets were more likely than those with fixed mindsets to give maladaptive person praise to their children (Gunderson et al., 2013). Although these findings might not necessarily translate to teachers, they indicate that more research is needed to clarify such relationships.

Clear relationships between teachers' achievement goals and classroom goal structures have been established. Previous studies found positive associations between teachers' self-reported performance-oriented practices and both performance-approach and -avoidance goals (Butler, 2012; Mascret et al., 2017), or only performance-approach goals (Retelsdorf et al., 2010). Teachers' performance-oriented practices were negatively associated with their goal to master tasks (Mascret et al., 2017), while their self-reported mastery-oriented practices were positively correlated with their mastery goal orientation (Butler, 2012; Retelsdorf et al., 2010). Student-perceived performance goal structures were also positively associated with teachers' performance-avoidance goals and negatively associated with teachers' mastery goals (Dresel et al., 2013). It has also been demonstrated that teachers with stronger performance goals tend to use social reference norms, that is, to judge a student's performance based on the performance of others (Retelsdorf & Günther, 2011). This suggests that they might be more likely to compare students in front of the class.

Little is known about relationships between the other motivational beliefs and evaluation practices of interest. Some studies showed a link between teachers' expectations and their feedback practices, including that teachers give more criticism and less praise to students they hold low expectations for (see Cooper, 1979). However, these studies do not indicate whether teachers gave effort or performance praise. Following Weiner's attribution theory (Weiner et al., 1971), teachers should reinforce attributes that they perceive to be responsible for students' performance because this could lead to performance improvements. That is, teachers who make effort attributions should be more likely to concentrate on students' effort when evaluating them.

2.2.3.2 Evaluation and students' motivational beliefs and achievement

Park et al. (2016) showed in a longitudinal study that the more performance-oriented practices teachers reported, the more their students endorsed fixed mindsets after controlling for initial beliefs. There was no significant relationship between mastery-oriented practices and student mindset, possibly due to a ceiling effect in teacher reports. Importantly, teachers' behaviours but not their mindsets predicted student mindset. In a different study, student-perceived mastery goal structure, assessed by the TARGET dimensions, did predict a growth mindset (Lüftenegger et al., 2017). Schmidt et al. (2015) showed that the initial success of mindset interventions was not sustained if students were subsequently taught by a teacher who emphasised performance compared to a teacher who made less use of performance-oriented practices. Competitive learning environments in physical education classes were also found to predict fixed mindsets amongst students (Ommundsen, 2001). Similarly, students were more likely to experience declines in growth mindset when they had teachers who used more social comparison (Dickhäuser et al., 2017). Moreover, children tend to adopt fixed mindsets after receiving person praise or praise that attributes performance to ability, whereas they tend to adopt growth mindsets after receiving effort praise or praise that attributes performance to effort from their parents or experimenters in a teacher role (Gunderson et al., 2013; Kamins & Dweck, 1999; Mueller & Dweck, 1998; Pomerantz & Kempner, 2013). Interestingly, it has been shown that the more students perceived process-oriented practices (e.g., the teacher accepts only full effort, the teacher wants students to think rather than just memorise), the less likely they were to think their teachers believed only some students could be good at maths and that their teachers preferred high performers (Ramirez et al., 2018).

Multiple studies demonstrated that students' perception of goal structures and mastery- or performance-oriented practices are precursors of corresponding goal orientations (e.g., Bardach et al., 2020; Bergsmann et al., 2013; Church et al., 2001; Meece et al., 2006; Midgley et al., 1995; Nolen, 2003; Lüftenegger et al., 2014; Schiefele & Schaffner, 2015; Urdan, 2004; Urdan & Schoenfelder, 2006; Wolters, 2004). For instance, Church et al. (2001) found that undergraduates' perceptions of their teachers emphasising the importance of grades and performance evaluations were negatively associated with mastery goals and positively associated with performance goals. Dresel et al. (2013) furthermore showed that student-perceived goal structures partly mediated the influence of teacher goals on student goals in maths. Teacher-reported mastery- and performance-oriented practices can also explain some of the variance in student goals (Anderman et al., 2001) but this relationship is usually weaker than relations with student-perceived practices (see Meece et al., 2006; Schiefele & Schaffner, 2015).

Research demonstrated that students who perceived an emphasis on mastery goals in their classrooms were more likely to attribute success to effort than students who perceived a performance emphasis (Ames & Archer, 1988). Moreover, perceptions of performance goals but not mastery goals were related to the tendency to attribute failure to a lack of ability. More evidence for a relationship between goal structure and attributions comes from a study that found that children who were praised for their intelligence rather than effort made low-ability instead of low-effort attributions for their poor performance on an experimental task (Mueller & Dweck, 1998). Similarly, young children who received more process praise from their parents (e.g., “good try”) were more likely to make effort attributions years later (Gunderson et al., 2013).

Evaluation practices appear to influence students’ self-concept as well. It has been found that students were more likely to experience declines in perceived maths ability when they had teachers who used more social comparison (Dickhäuser et al., 2017). Similarly, teachers’ use of an individual frame of reference to evaluate students is positively associated with students’ self-concept (Lüdtke & Köller, 2002). Student-perceived mastery goal structures also tend to positively correlate with their competence beliefs (for a review, see Givens Rolland, 2012). Some studies found a negative relationship between performance goal structures and students’ perceived competence, while others found no relationship (see Givens Rolland, 2012). A study with undergraduate students found that those receiving person praise (e.g., praise for natural talent) rather than process praise (e.g., praise for effort) after a successful task reported lower perceived competence following failure (Haimovitz & Corpus, 2011). Similar detrimental effects of person praise on perceived performance were found for British schoolchildren aged 9–11 years as well as university students (Skipper & Douglas, 2012).

There is some indication that performance-oriented practices may negatively affect students’ achievement. A study with university students showed that less emphasis on learning and development was associated with lower performance, especially for minority students (Canning et al., 2019). Another study found a negative relationship between performance-oriented practices and achievement, but the results were non-significant (Park et al., 2016), and one study found a weak, positive relationship between mastery-oriented practices and maths achievement (Wolters, 2004).

2.2.3.3 Brief summary on evaluation and mindset

There is much evidence suggesting that teachers with fixed mindsets engage in more performance-oriented practices compared to those with growth mindsets using teacher reports (Miele et al., 2019; Park et al., 2016), student reports (Canning et al., 2019; Dresel et al., 2013), and observations (Stipek et al., 2001). However, Sun (2019) observed that teachers with growth mindsets did not consistently avoid using comparative structures, which was unfortunately not compared to fixed mindset teachers. Multiple studies showed that performance-oriented practices foster students' fixed mindsets (Dickhäuser et al., 2017; Ommundsen, 2001; Park et al., 2016; Schmidt et al., 2015), whereas mastery goal structures predict growth mindsets (Lüftenegger et al., 2017). Overall, previous studies provided ample evidence for relationships between performance-oriented practices and the motivational constructs of interest. However, given that many of the reviewed studies examined goal structures and performance-oriented practices, which are rather broad dimensions, it is difficult to gauge in how far these findings translate to performance over process recognition and social comparison specifically.

2.2.4 Little autonomy

The little autonomy dimension includes two aspects: *little choice* (e.g., students cannot choose their tasks or work in their own way) and *little promotion of self-regulation* (e.g., students do not set their own learning goals or evaluate their own progress). The study of autonomy-supportive climates is mostly associated with literature on self-determination theory (Deci & Ryan, 2000). Yet, this dimension is included in the present study because important theoretical and empirical links between autonomy support and mindset have been identified as this section will show. Promoting student autonomy through choice and self-regulation is largely incompatible with controlling practices, including that the teacher regulates which students work on what tasks, with whom and when, as well as what strategies they use to solve problems (Bohlmann & Weinstein, 2013; Stipek et al., 2001). Nevertheless, autonomy support and control are not direct opposites, so that the absence of autonomy support does not automatically imply engagement in controlling practices (see Bartholomew et al., 2018). I decided to focus on the absence of autonomy support rather than controlling practices because previous research, as reviewed below, indicated that teachers with fixed mindsets are less likely to promote autonomy, while I found no previous evidence showing these teachers are more controlling. As such, fixed mindset teachers may simply not actively support student autonomy rather than that they actively frustrate it.

2.2.4.1 Teachers' motivational beliefs and little autonomy

Teachers with a fixed mindset are less likely to report that they establish an autonomy-supportive climate than growth mindset teachers (Leroy et al., 2007). Stipek et al. (2001) found a negative but non-significant correlation between teachers' fixed mindsets and their autonomy support as rated by observers (r= -.43). They proposed that teachers with fixed mindsets may believe that less-able students cannot make productive use of their autonomy. Similar relationships were identified in studies with parents, although these looked at controlling behaviour rather than the absence of autonomy support. For example, mothers who were induced to hold fixed mindsets were more likely than those with growth mindsets to be controlling with their children while working on experimental tasks (Moorman & Pomerantz, 2010). Likewise, parental growth mindset was found to be negatively related to their controlling behaviour (Matthes & Stoeger, 2018).

Turning to other motivational beliefs, a study showed that teachers who created performance-oriented classrooms tended to insist that students correctly followed prescribed procedures (Patrick et al., 2001). Teachers who were oriented towards extrinsic goals similar to performance goals also tended to use less autonomy support and more control (Jang, 2017). However, another study found no association between teachers' mastery goal and self-reported autonomy support (George & Richardson, 2019). Teachers' expectations may be linked to autonomy support because teachers with low expectations are less likely than those with high expectations to think that students should be given ownership over their learning and to allow students to choose their activities (Rubie, 2003). Yet, a different study found no significant relationship between teachers' expectations and student-perceived autonomy support in physical education classes (Trouilloud et al., 2006). I found no studies that examined relationships between teachers' attributions and autonomy support. It may be that teachers who attribute students' achievement mostly to ability rather than effort are less likely to provide choice because they may tend to assign students only those tasks that they think students can solve based on their perceived ability.

2.2.4.2 Little autonomy and students' motivational beliefs and achievement

Ommundsen (2001) showed that opportunities for choice predicted students' growth mindsets in physical education classes. He argued that these opportunities might cause students to feel greater control over the learning process, that their current ability must not be permanent, and that working hard can be fruitful. In addition, student-perceived autonomy support was related to adopting mastery- and normative performance-approach goals, whereas perceived teacher control was related to adopting performance-avoidance goals (Shih, 2013). Similarly, student-perceived autonomy support was found to be positively correlated with mastery goals and negatively correlated with appearance performance-approach and -avoidance goals, whereas autonomy suppression correlated negatively with mastery goals and positively with appearance performance-approach and -avoidance goals (Madjar et al., 2013). Madjar et al. (2013) emphasised the importance of employing a longitudinal design to gain greater insights into a potential causal relationship. One study doing just that did not find that autonomy support related to students' mastery goals (Ruzek & Schenke, 2018). This was further examined in the present study. Lastly, a study found that students perceived higher competence when they experienced more autonomy support in physical education classes (Trouilloud et al., 2006). This was also shown for academic subjects (Deci et al., 1981).

Teachers' autonomy support has been demonstrated to be adaptive for students' educational cognitions and behaviours, such as engagement, quality of learning, preference for challenge, and intrinsic motivation, as well as higher achievement (for a review, see Su & Reeve, 2011).

2.2.4.3 Brief summary on little autonomy and mindset

Previous studies indicated that teachers with a fixed mindset are less likely to engage in or value autonomy-supportive practices (Leroy et al., 2007; Stipek et al., 2001), potentially because these teachers may believe that less-able students cannot make good use of autonomy (Stipek et al., 2001). Additionally, one study showed that having choice predicted students' growth mindsets, which was proposed to be due to a greater sense of control over the learning process (Ommundsen, 2001). Clearly, more evidence is needed to substantiate the results of these few studies.

2.2.5 Differential value

The *differential value* dimension relates to teachers showing that they value high-performing students more than low-performing students. It includes aspects such as listening more carefully to high performers and asking them more frequently to give answers in front of the class, thereby not encouraging all students to participate equally.

2.2.5.1 Teachers' motivational beliefs and differential value

Decades ago, Marshall and Weinstein (1984) proposed that teachers with fixed mindsets may be more likely than those with growth mindsets to "place students in a permanent hierarchy according to their expectations and to treat them differently" (p. 305). It has been suggested that teachers endorsing fixed mindsets may invest more resources in high-performing students because they may think that helping a student with a stable, low ability is fruitless (Leroy et al., 2007; Rissanen et al., 2018). This also points towards a potential influence of attributing performance to a stable ability on differential value. In contrast, teachers with a growth mindset might treat students more equitably or be even more motivated to teach lower-performing students because they believe all students can develop their ability if they are challenged enough (Rissanen et al., 2018). Similarly, it has been observed that teachers who were perceived to foster mastery goals were more likely to require all students to get involved in the lesson, whereas teachers who were perceived to foster performance goals did not ensure participation by all (Patrick et al., 2001). Cooper and Good (1983) also proposed that teachers may discourage low-expectation students from orally participating in the classroom. In addition, teachers who were perceived to foster performance goals were more likely to use threats and punishments, such as assigning extra homework, compared to mastery-fostering teachers, who showed more warmth towards students (Patrick et al., 2001). Teachers also tend to show more negative affect and less warmth towards students for whom they have lower expectations (see Babad, 1993; Rubie-Davis, 2009).

2.2.5.2 Differential value and students' motivational beliefs and achievement

I found no studies that examined links between teachers' differential value and students' mindsets, goal orientations, or attributions. Most research looking at differential value or treatment focused on expectancy effects, and thus students' self-concept. For instance, teachers tend to treat students for whom they have lower expectations with less warmth and give them fewer opportunities to respond in the class, which, in turn, are both associated with lower student performance, presumably through expectation effects (for a review, see Babad, 1993). Similarly, Bohlmann and Weinstein (2013) proposed that teachers' judgements of students' ability become particularly apparent when teachers are more accepting or warm towards some students. Students who feel less valued may assume that teachers expect little of them and that working hard will be futile. In line with this, student-perceived emotional support, such as feeling safe, was shown to be positively related to their self-concept (Skaalvik & Skaalvik, 2013).

Differential value may be related to students' achievement because teacher warmth (e.g., teachers listen to students and do not put them down) was found to be positively related to student achievement in a range of subjects, including maths (Voelkl, 1995). This was fully explained through relationships between warmth and student participation in class. Accordingly, if teachers do not encourage all students to participate in class, the achievement of those who are not taking part is likely to suffer.

2.2.5.3 Brief summary on differential value and mindset

Teachers' fixed mindsets have theoretically been linked to their differential valuing of low- and high-performing students (Leroy et al., 2007; Marshall & Weinstein, 1984; Rissanen et al., 2018). However, no empirical evidence for this relationship has been provided yet. Similarly, I found no studies examining teachers' differential value in relation to students' mindsets. Thus, it becomes very clear that more research is needed to understand how differential value is related to teachers' and students' mindsets.

2.2.6 Little promotion of risk-taking

This dimension relates to teachers' *little encouragement of help-seeking* (e.g., the teacher does not answer questions carefully and thoroughly) and their *unconstructive handling of mistakes* (e.g., the teacher is critical of mistakes). The extent to which students feel safe to take intellectual risks, that is, to risk revealing they struggle with schoolwork by making mistakes or asking for help, may impact their motivation (Stipek et al., 1998). Multiple scholars argued that emphasising that growth often requires struggling and making mistakes can foster growth mindsets in students (Anderson et al., 2018; Boaler, 2016; Haimovitz & Dweck, 2017).

2.2.6.1 Teachers' motivational beliefs and little promotion of risk-taking

Teachers with fixed mindsets may be less likely to handle mistakes adaptively than teachers with growth mindsets because the former tend to view mistakes as debilitating and thus as something that must be avoided (Dweck, 1999). They may also regard asking questions as a sign of inadequate ability (Karabenick, 1998). Hence, fixed mindset teachers may tend to discourage students from making mistakes and asking questions, and if students do err or seek help, to respond with criticism or ignorance. In contrast, growth mindset teachers might convey that it is ok to ask for help and make mistakes. Stipek et al. (2001) indeed observed that teachers with fixed mindsets were less likely to handle mistakes adaptively, but this correlation was non-significant. Moreover, Sun (2019) found that growth mindset teachers tended to stress the importance of mistakes but that they did not consistently engage with mistakes in ways that improve students' conceptual understanding. Growth mindset teachers also tended to offer opportunities for extra help. Due to the lack of a comparison group, it is difficult to draw clear conclusions from this. Related evidence comes from a study showing that mothers of primary school students were more likely to regard failure as debilitating when they endorsed fixed rather than growth mindsets (Haimovitz & Dweck, 2016). Moreover, mothers who were induced to have a fixed mindset were more likely to react unconstructively to their children's mistakes, such as providing correct answers themselves or stating that the child did a mistake without explaining why, during experimental tasks (Moorman & Pomerantz, 2010).

Moving on to the other motivational beliefs, a study showed that students were more likely to perceive mastery-oriented teachers to encourage help-seeking, whereas they perceived teachers with performance-avoidance goals to inhibit questioning and to convey that help-seeking implies a lack of ability (Butler & Shibaz, 2008). In addition, an observational study showed that performance-oriented teachers were more concerned about right answers while mastery-oriented teachers repeatedly told students that mistakes belonged to the natural learning process (Patrick et al., 2001). Similarly, teachers are more likely to criticise wrong answers of students for whom they have low expectations, while they are more likely to rephrase questions for high-expectation students (Brophy & Good, 1970). I found no studies that examined associations between teachers' attributions and promotion of risk-taking.

2.2.6.2 Little promotion of risk-taking and students' motivational beliefs and achievement

Previous research suggests a potential relationship between unconstructive handling of mistakes and students' fixed mindset. For instance, students who reported that they have to worry about making mistakes during physical education classes tended to adopt fixed mindsets (Ommundsen, 2001). Similarly, mothers who regarded mistakes as debilitating were more likely to have children with fixed mindsets (Haimovitz & Dweck, 2016). Children's perceptions of their mothers' views towards failure were related to children's mindsets, and this relationship was fully mediated by children-perceived parental goal orientations. I found only one study that examined relationships between encouragement of help-seeking and student mindset. Surprisingly, that study found that university instructors' help messages, including encouragement to ask for individual help, were related to the adoption of stronger fixed mindsets amongst students (Barger, 2019). The author suggested that students may interpret help offering as a sign of not being able to learn to do the tasks on their own. This does not seem entirely plausible because these were general messages about asking for help if needed rather than giving unsolicited help to certain students. This is further investigated in the present study.

I could not find studies investigating relationships between little promotion of risk-taking and students' achievement goals. However, it seems reasonable that feeling as though it is not ok to make mistakes and asking for help may foster performance goals. This is because being discouraged from asking questions may convey to students that understanding is less important than appearing smart (for a review of relationships between student goal orientation and help-seeking, see Butler, 2006). Similarly, if making mistakes is regarded as something that should be avoided, students may begin to believe that performance outcomes are more important than trying. If teachers do not convey that students should seek help or that mistakes can benefit learning, students may also think that effort is fruitless and, thus, that their achievement is largely determined by their ability. Lastly, Bohlmann and Weinstein (2013) proposed that teachers' judgements of students' ability become particularly apparent when teachers are dismissive or ignore students when they give wrong answers. This is then likely to affect students' self-concept.

Little encouragement of help-seeking is likely to reduce students' achievement because the more students ask for help, the more progress they tend to make (Schenke et al., 2015). Importantly, Schenke et al. found that student-perceived mastery goal structure and teacher emotional support predicted their help-seeking. Another study showed that students who thought that their teachers had positive attitudes towards mistakes and handled them constructively, such as explaining mistakes and giving the chance to try again, had higher achievement (Käfer et al., 2019). This seems plausible because students who feel safe to attempt tasks although they might make mistakes have more learning opportunities. Additionally, discussing mistakes presumably leads to greater understanding and, thus, achievement gains. Moreover, one study demonstrated that an adaptive "error climate" in maths lessons predicted more investment of effort, which was mediated by students' reactions to mistakes (Steuer et al., 2013). More investment of effort is likely to benefit achievement over time.

2.2.6.3 Brief summary on little promotion of risk-taking and mindset

There are theoretical links between having a fixed mindset and regarding mistakes as something that should be avoided as well as considering asking questions as a sign of inadequate ability (Dweck, 1999; Karabenick, 1998). But there is little empirical evidence showing that teachers with fixed mindsets handle students' mistakes and help-seeking unconstructively. A qualitative study indicated that growth mindset teachers tend to adaptively handle mistakes and help-seeking, but the study did not include a comparison group of fixed mindset teachers (Sun, 2019). Some related evidence comes from studies looking at parents' mindsets and their attitudes towards their children's mistakes (Haimovitz & Dweck, 2016; Moorman & Pomerantz, 2010), but whether these findings translate to teaching practices is not certain. Likewise, research on relationships between teachers' promotion of risk-taking and students' mindsets is sparse, which demonstrates the need for further work. Only one study found that having to worry about making mistakes fostered fixed mindsets (Ommundsen, 2001). Some related evidence comes from studies with parents and children (Haimovitz & Dweck, 2016). One study even found that university instructors' encouragement to ask for individual help fostered fixed mindsets (Barger, 2019), which contradicts theory. There is thus clearly need for more research.

2.3 Objective and research questions

2.3.1 Need for study and objective

As reviewed in this chapter, students' belief that their ability is relatively fixed as opposed to malleable is related to a range of other unhelpful motivational beliefs, including an endorsement of performance- rather than mastery-goals, attributing performance to ability rather than effort, lower self-concept, and even lower academic achievement. This is especially problematic in maths, where fixed mindsets are particularly prevalent. Although the importance of students' mindsets in educational settings has been established, research examining why some students develop a fixed mindset is scarce.

Although there is some prior evidence that FAPs are related to students' motivational beliefs and achievement and that teachers' beliefs are linked to their engagement in FAPs, the findings are not conclusive but can rather be viewed as a starting point. Some authors only proposed certain relationships between practices and student mindset on theoretical grounds, such as between ability grouping and fixed mindset, but did not test this directly (Boaler, 2013; Boaler et al., 2000; Marks, 2013). Moreover, I presented findings from research looking at a physical education context (Ommundsen, 2001; Trouilloud et al., 2006), which may differ from maths, as well as from research looking at interactions between mothers and children (Haimovitz & Dweck, 2016) or researchers and children (Kamins & Dweck, 1999; Mueller & Dweck, 1998) rather than teachers and students. There is also a common reliance on hypothetical scenarios (Lee, 1996; Rattan et al., 2012), teacher-reports alone (Leroy et al., 2007; Mascret et al., 2017; Park et al., 2016; Retelsdorf et al., 2010; Shim et al., 2013), student-reports alone (Butler & Shibaz, 2008; Dresel et al., 2013), or cross-sectional designs (for exceptions, see Park et al., 2016; Retelsdorf et al., 2010; Butler & Shibaz, 2008).

Previous studies have investigated the relationships of interest in a rather piecemeal fashion by considering only a few isolated practices rather than a set of relevant, interrelated practices, which might overestimate the importance of a given practice. Only when examining a range of practices simultaneously can we estimate the relative contribution of each practice to shaping fixed mindsets. In addition, it is important to focus on practices that are directly relevant to mindset. Some studies examined the impact of classroom goal structures or performance-oriented practices, which are based on achievement goal orientation, on mindset (Lüftenegger et al., 2017; Park et al., 2016). Although achievement goals are clearly related to mindset, they are theoretically distinct, so that practices that are important when taking a mindset-lens might have been overlooked.

To advance this line of research, the present study first aimed to develop a comprehensive conceptualisation of fixed-ability practices as well as a measurement tool for students and teachers to assess FAPs. It then aimed to examine the dimensionality of these practices. As a large sample is needed to examine the factor structure of FAPs, the dimensionality was only based on student reports. Next, it aimed to investigate whether teachers' own motivational beliefs predict their use of student-perceived and/or teacher-reported FAPs. Moreover, the study aimed to examine the longitudinal effects of student-perceived and teacher-reported FAPs on students' fixed mindsets and associated motivational beliefs as well as achievement in a maths context. Lastly, to understand the mechanisms behind these relationships, the study qualitatively explored teachers' reasons for engaging in FAPs and students' perceptions of FAPs in relation to their beliefs.

Identifying teaching practices that foster fixed versus growth mindsets is crucial for researchers' and practitioners' efforts to foster adaptive beliefs amongst students because these practices provide important targets for interventions. Addressing the antecedents of students' fixed mindsets might be more feasible than intervening after maladaptive beliefs have been established; While student interventions can successfully induce growth mindsets (e.g., Blackwell et al., 2007; Good et al., 2003), sustained success is likely to depend on the teaching practices that students subsequently experience (Schmidt et al., 2015). This study can increase practitioners' awareness of how certain behaviours may affect their students and provide them with practical advice on how they can avoid fostering maladaptive beliefs. It is thus also important to understand why teachers engage in FAPs – introducing change to practices without addressing the underlying beliefs might be ineffective (Cross, 2009; Stipek et al., 2001). Likewise, solely informing teachers about the importance of growth mindset may also not be sufficient as it cannot be assumed that they directly pass on these beliefs to students (see Haimovitz & Dweck, 2017).

Before moving on to the research questions, I briefly want to acknowledge that other factors than examined in the present study can influence the constructs of interest. First, as already reviewed above, not only teaching practices but also parents' beliefs and practices play an important role in the development of students' motivational beliefs. It also seems plausible that the wider school environment, independent of the practices of individual teachers, affects motivational beliefs, such as the use of league tables. In addition, researchers have recently shown that the mindsets of peers can influence students' mindsets (King, 2019). Similarly, there are other potential factors besides motivational beliefs that could influence teachers' engagement in FAPs. Besides the motivational beliefs examined here, teaching practices can be influenced by other beliefs relevant to the teaching context as well as teachers' pedagogical and content knowledge. This includes, for instance, beliefs about the school climate or culture, beliefs about the subject, beliefs or knowledge about specific teaching practices, and even beliefs about the general teaching approach, such as constructivism and transmission (see Fives & Buehl, 2012). For this study, teachers' mathematical knowledge would have been especially relevant as it has been found to influence the teaching practices in which they engage during maths lessons (Jacobson & Izsák, 2015). Moreover, outside pressures such as the accountability culture in the UK or school policies might drive teachers to emphasise performance more than they would like to. Accordingly, this study does not claim to take all or even most factors into account, which was, of course, impossible with the time and resource constraints. Nevertheless, it did aim to identify highly relevant aspects that lend themselves as targets for future interventions.

2.3.2 Research questions

Based on the existing literature and the need for further research as just described, the present study aims to answer the following research questions:

1. What are fixed-ability practices? More specifically, what are the individual dimensions of fixed-ability practices, at least as perceived by students, and is there an overarching fixed-ability dimension?

2. Do teachers' motivational beliefs predict their engagement in student-perceived and/or teacher-reported fixed-ability practices? If so, which beliefs predict what practices?

3. Do student-perceived and/or teacher-reported fixed-ability practices predict students' motivational beliefs and achievement? If so, which practices predict what beliefs and which practices predict achievement?

4. How do teachers justify their engagement in fixed-ability practices, and are these justifications related to their motivational beliefs?

5. How do students perceive teachers' fixed-ability practices, and what messages do they draw from these in relation to their motivational beliefs?

Chapter 3
Research Methodology

This chapter details the methodology of the present study. In the first section (Section 3.1), it describes and justifies the research design that was chosen to answer the research questions (see Table 3.1 for an overview of the research questions, the corresponding data sources, and planned analytical approach). It then describes the methods that were used (Section 3.2). This is followed by a description of the sampling strategy, the setting, and the student and teacher sample (Section 3.3). Then, the specific quantitative (Section 3.4) and qualitative (Section 3.5) research instruments, as well as the type of achievement data (Section 3.6) are described. Subsequently, the procedures, including two pilot studies, are presented (Section 3.7). The concluding section (Section 3.8) addresses important ethical considerations. Please note that the analytical strategies are presented in later chapters.

3.1 Research design and theoretical perspective

This study used a longitudinal mixed-methods design as shown in Figure 3.1. Teachers' and students' mindsets and motivational beliefs were quantitatively assessed through questionnaires at the beginning of the school year (Time 1), and students' beliefs were assessed a second time towards the end of the year (Time 3). Fixed-ability practices (FAPs) were assessed during the middle of the year (Time 2) using teacher- and student-reports. Teachers' beliefs and FAPs were assessed only once as these constructs are fairly stable over time (for teacher beliefs including mindset, see Pomerantz & Saxon, 2001; Robins & Pals, 2002; Stipek et al., 2001; for teaching practices similar to those considered in the present study, see Anderman et al., 2002). Classroom observations and subsequent video-stimulated recall interviews were conducted with a subset of teachers and some of their students during the middle of the school year (Time 2) to encourage reflections on FAPs. Lastly, information on students' maths achievement at the end of Year 6[2] and Year 7 was obtained from school records. The study was conducted with secondary school maths teachers and their Year 7 students.

[2] This is the last year of primary school in England.

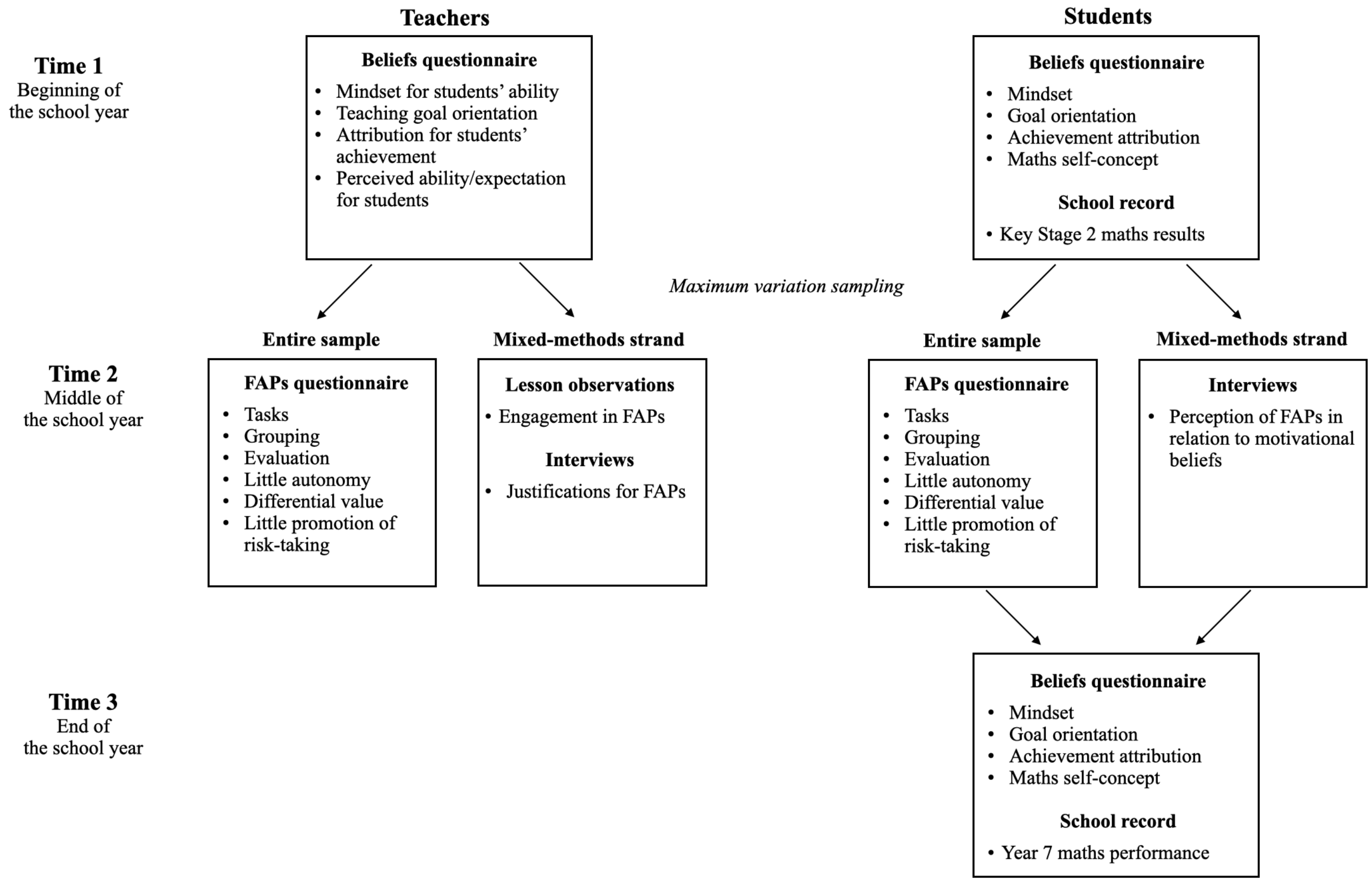
Table 3.1

Overview of the Research Questions

Research question	Data source	Planned data analysis	Results chapter
1. What are fixed-ability practices? More specifically, what are the individual dimensions of fixed-ability practices, at least as perceived by students, and is there an overarching fixed-ability dimension?	Student questionnaires	Exploratory and confirmatory two-level factor analysis	5
2. Do teachers' motivational beliefs predict their engagement in student-perceived and/or teacher-reported fixed-ability practices? If so, which beliefs predict what practices?	Student questionnaires Teacher questionnaires	Correlations and multivariate regression	7
3. Do student-perceived and/or teacher-reported fixed-ability practices predict students' motivational beliefs and achievement? If so, which practices predict what beliefs and which practices predict achievement?	Student questionnaires Teacher questionnaires	Bayesian multivariate cross-classified multiple membership random effects model	7
4. How do teachers justify their engagement in fixed-ability practices, and are these justifications related to their own motivational beliefs?	Teacher interviews	Thematic analysis	9
5. How do students perceive teachers' fixed-ability practices, and what messages do they draw from these in relation to their motivational beliefs?	Student interviews	Thematic analysis	10

Figure 3.1

Graphical Representation of the Study Design



Note. FAPs = Fixed-ability practices.

I used a mixed-methods approach because neither quantitative nor qualitative methods alone were sufficient to answer the research questions and because employing both enabled a better understanding of the research problem (Biddle & Schafft, 2015; Creswell & Plano-Clark, 2011). Accordingly, I followed a pragmatic approach to ensure that the methods were suitable to answer the research questions (Teddlie & Tashakkori, 2009). In this study, quantitative methods were best suited to identify and describe patterns or relationships, and the magnitude of such relationships, between teachers' motivational beliefs and their teaching practices as well as between teaching practices and students' motivational beliefs and achievement across a large number of participants. Qualitative methods were used to enrich and interpret these more descriptive relationships based on the meaning-making processes of selected teachers and students. The advantages of each specific method are further described in the next section. I used a longitudinal design because examining whether FAPs predict students' motivational beliefs and achievement is only possible when initial levels of beliefs and achievement are controlled for. Such a design can help to establish the direction and magnitude of the examined relationships, whereas a cross-sectional design can only provide information of what is present at a given time (Gravetter & Forzano, 2012).

3.2 Research methods

This section describes the methods of this study and provides the rationale for their use. The specific research instruments are described later (see Sections 3.4 and 3.5).

3.2.1 Assessing motivational beliefs

Teachers' and students' mindsets and motivational beliefs were assessed through questionnaires. Looking at the education and psychology literature, this is probably the most prevalent approach to assessing individuals' motivation. The use of questionnaires is based on the assumptions that participants understand the questions as intended, have access to the answers, consider their answers appropriately, and are willing to share their views honestly (Cohen, 2012). To meet these assumptions, instruments were selected that demonstrated validity and reliability in previous studies whenever available. Moreover, all instruments were piloted and adapted if necessary. This was particularly important as the students were relatively young and may have had difficulties understanding certain words or sentences. To reduce a potential social desirability bias, it was emphasised that all responses would be treated confidentially, and that data would be anonymised following the completion of data collection (Cohen, 2012).

3.2.2 Assessing fixed-ability practices

Researchers examining the effects of teaching on student outcomes have highlighted the importance of using multiple sources of information to gain a fuller picture of classroom practices (Babad, 1993; Raudenbush, 2005; Seidel & Shavelson, 2007). Hence, teacher and student reports were employed to assess FAPs. This is not just a matter of using different sources to assess the same practices (i.e., triangulation) because teachers and students may have access to different aspects of teaching (Seidel & Shavelson, 2007). Each method has its strengths and limitations. While teachers can be a valuable source of information because they know the purpose of their practices (Wolters et al., 2010), they may answer in socially desirable ways or report practices they aim to employ rather than those they actually do employ (Fang, 1996). Accordingly, teacher-reported practices are usually more favourable than those reported by students (Butler, 2012; Deemer, 2004; Schiefele & Schaffner, 2015).

Considering the students' perspective is essential as they may perceive practices in different ways than teachers intended (Babad, 1990; Deemer, 2004). From a social-constructivist point of view, these differences in perceptions might be most relevant since a given student's beliefs are more affected by their own perception or interpretation of teaching practices than by any objective indicator (Lüdtke et al., 2009; Miller & Murdock, 2007). However, the validity of student reports has also been questioned because they may be influenced by their personal experiences, background, and personality characteristics, such as their competitiveness (Wolters et al., 2010). Overall though, using student reports to assess the learning environment is advantageous (Lüdtke et al., 2009). I should point out that student ratings can and were used at two different levels of analysis. First, individual student ratings of FAPs can be interpreted as students' individual perception of practices and, second, aggregated student responses can be interpreted as the collective perception of these practices (Lüdtke et al., 2009).

Note that I also observed maths lessons of selected teachers, which can also be an effective tool for assessing teaching practices (Wolters et al., 2010). However, observations were not used to assess FAPs because this went beyond the scope of this PhD given my time and resource constraints. Instead, observations were used to subsequently discuss teaching practices in teacher and student interviews, which is described next.

3.2.3 Exploring why teachers use fixed-ability practices and how these affect students' motivational beliefs

In addition to identifying whether FAPs predict students' motivational beliefs, this study also sought to gain greater understanding of why such relationships might exist as well as why teachers may use these practices. For this purpose, observations were conducted during maths lessons of four selected teachers, and, subsequently, individual and group video-stimulated recall interviews were conducted with these teachers and subsamples of their students, respectively. Video-stimulated recall interviews are an effective tool for stimulating deliberate and critical reflection of teaching practices in maths with teachers (Muir, 2010) as well as students (Tanner & Jones, 2007). Video segments can help interviewees to articulate their thinking through a clear focus and context for the interview questions (Powell, 2005). More generally, interviews allow for a direct and instant communication between the researcher and participant and are thus particularly useful to gain an understanding of individuals' meaning-making processes (Cohen, 2012; Gillham, 2000). Participants can talk freely and go into more depth regarding issues that are most relevant to them (Mears, 2009), without being limited to a narrow set of questions or response categories. Despite these advantages, there are important challenges when conducting interviews. As the researcher, I inevitably brought my preconceived ideas and assumptions about the relationships between practices and beliefs to the interviews. Hence, I tried to be especially careful not to steer participants in any particular direction and to avoid the creation of demand characteristics by asking open questions and avoiding leading questions (Gomm, 2004). It was also essential to establish good rapport with participants because interviews rely on the collaboration between the researcher and participant (DiCicco-Bloom & Crabtree, 2006; Mears, 2009).

3.3 Sampling, setting, and participants

3.3.1 Sampling

A sequential mixed-methods sampling approach was taken (Teddlie & Yu, 2007). First, maths teachers and students from 37 Year 7 classrooms in seven secondary schools were recruited for the overall sample using a convenience approach. Second, from this overall sample, four teachers and some of their students were selected for the qualitative part of the study based on maximum variation sampling. Year 7 was chosen as it is the first year of secondary education in England and thus marks the transitional period from primary to secondary school. This period is highly relevant for the formation of students' ability beliefs, presumably due to drastic changes in the school environment, characterised by an increased emphasis on ability and performance (Eccles et al., 1993; Eccles & Roeser, 2009). Another consideration was that participating teachers and students did not know each other from previous years so that the relationships between fixed-ability practices and students' beliefs and achievement could be examined from the beginning onwards.

3.3.1.1 The overall sample

It would have been desirable to use probability sampling to randomly select secondary schools in England as this would have provided greater external validity for the findings. However, this was not feasible due to practical considerations, which is why only schools that could be reached relatively easily were invited to participate. Study invitations were sent to headteachers or heads of maths of state-funded co-educational and single-sex secondary schools in the East of England. Moreover, a call for participants was posted on local Maths Hub websites[3] and distributed through their email lists. Personal networks at the Faculty of Education were also utilised to identify and contact potential secondary schools.

Students in the sample belonged to different maths classes and teachers, and some students even belonged to multiple teachers and classes over the course of the study. This is explained in more detail in the next chapter (see Section 4.1.2). Due to the complex nature of the quantitative analyses that controlled for students being nested in multiple teachers and classrooms, it was not possible to perform power calculations before data collection. Instead, the literature on multilevel modelling was used to gauge the necessary minimum sample size. Generally, the sample size at the highest level (here classroom and teacher because schools were not modelled as a level as will be explained in Section 4.1.2) is the main limiting factor. Rules of thumb vary greatly but a simulation study showed that a group number of as low as 20 may be sufficient for two-level multiple regressions (Maas & Hox, 2005). It demonstrated that regressions with such small sample sizes yielded accurate estimations of fixed regression coefficients and random variance components, as well as standard errors of fixed regression coefficients, using a medium effect size (.30). However, standard errors of the level-two variances were estimated as too small. Small sample sizes in two-level designs are a frequent issue in social and educational research. With this in mind, practical considerations were at the forefront; the aim was to recruit as many schools and classes as possible before the start of

[3] Maths Hubs belong to a network of schools and teachers working to lead improvement in maths education in England.

the school year during which data was collected. In total, 37 classes from seven schools were recruited.

3.3.1.2 The sample for the mixed-methods strand

Following the first data collection phase, a subsample of the original sample was selected using maximum variation sampling. This subsample participated in the observations and interviews in addition to completing the questionnaires. The mixed-methods strand consisted of four teachers and some of their students. Note that only teachers from co-educational schools were considered to aid comparability, and one school was excluded due to its location. Maximum variation sampling involves identifying key dimensions on which participants vary and then selecting participants who vary from each other as much as possible (Suri, 2011). In this study, the selection was based on teachers' and students' motivational beliefs reported at Time 1 as well as students' current performance as judged by teachers. The rationale for this sampling strategy was that teachers with the greatest differences in their beliefs presumably differ in their practices, which would thus enable me to explore various degrees of FAPs. Similarly, interviews about students' perceptions of these practices are presumably most informative with students who have different motivational beliefs and performance levels. Because mindset is at the centre of the meaning system comprised of the motivational beliefs of interest, as explained in the introduction, the sample selection was mostly based on mindset. Unfortunately, one of the initially selected teachers was not willing to participate in observations, which is why he was replaced by the most appropriate alternative. Moreover, while I selected four boys and four girls from each class, not all of these were willing to be interviewed and were thus replaced. In addition, some students spontaneously changed their minds before the interviews, so that group sizes varied. The four selected teachers are further described below (see Section 3.3.3.2).

3.3.2 Setting

Seven secondary comprehensive schools from the East of England participated in the study. Five schools were co-educational, one was a girls-only school, and another a boys-only school. An overview of these schools with descriptive characteristics is provided in Table 3.2. Characteristics were obtained from OFSTED reports that were available online. The teachers and students from the mixed-methods strand belonged to School 1 (Meadows School[4]) and School 7 (Woodpark School).

[4] Schools are referred to by pseudonyms.

Table 3.2

School Characteristics

	School type	Sex	Student ethnicity	Student premium eligibility	Progress 8 score	Participating classes (N)	Participating teachers (N)	Maths ability grouping
School 1 (Meadows School)	Free School	Mixed	High ethnic minority proportion	Well above average	Well above average	5	5	All classes are streamed based on combined performance in multiple subjects
School 2	Academy converter	Mixed	Large White British majority	Around the average	Below average	11	6	Only the top and bottom are set, the rest is mixed
School 3	Academy converter	Mixed	White British majority	Well below average	Well above average	2	1	All classes are set
School 4	Academy converter	Mixed	White British majority	Below average	Well above average	5	8	Only the bottom is set
School 5	Voluntary aided	Boys	Majority is of minority ethnic heritage	Well above average	Well above average	7	3	All classes are set
School 6	Academy converter	Girls	White British majority	Slightly below average	Well above average	5	6	No setting in Year 7
School 7 (Woodpark School)	Academy converter	Mixed	Large White British majority	Below the average	Below average	2	2	All classes are set

Note. Information on student ethnicity and eligibility for student premium was taken from the most recent OFSTED reports available at the beginning of the study. The Progress 8 score indicates the progress of students who graduated at the end of the academic year during which data was collected between the end of KS2 and the end of KS4, compared to students across England who had comparable results at the end of KS2.

3.3.3 Participants

3.3.3.1 Participants of the overall sample

The teacher sample consisted of 31 maths teachers (13=male, 18=female). Many teachers taught more than one participating class and many classes were simultaneously taught by multiple teachers or changed teachers over the course of the year. Notably, some classes were taught by additional teachers who did not actively take part in this study. Relevant demographic information of teachers can be found in Table 3.3. The total student sample consisted of 927 students from 37 Year 7 classes (typically aged 11-12 years; 478=male, 448=female, 1=female sex but identified as neither male nor female).

Table 3.3

Demographic Information of the Teacher Sample

Characteristics	Sample distribution % (N)	*M (SD)*
Education		-
Bachelor's Degree	26.7 (8)	-
PGCE	36.7 (11)	-
Master's Degree	36.7 (11)	-
Doctoral Degree	0 (0)	-
Years of teaching	-	7.17 (7.47)
Years of teaching maths	-	5.20 (4.06)

Please note that not all participants completed all questionnaires as some teachers or students either joined or left during the year or changed their minds regarding participation. The greatest attrition was caused by students being absent during collection days rather than a change of mind. The first questionnaire was completed by 30 teachers and 882 students. Of the student sample, eight were excluded for giving no or invalid responses (e.g., selecting two answer options) to more than 25% of the items. An additional student was excluded because of careless responding as indicated by the longstring method, which is further described in Section 4.1.3. The second questionnaire was completed by 28 teachers and 840 students. Eight students were excluded for giving no or invalid responses to more than 25% of the items, and one student was excluded because of careless responding. The third questionnaire was completed by 816 students. A large number of participants was unfortunately not present because of extracurricular activities on the day of data collection in two schools. Seven participants were excluded as they gave no or invalid responses to more than 25% of the items. Two additional participants were excluded because of careless responding. Please note that achievement data was not available for all students.

3.3.3.2 Participants of the mixed-methods strand

Observations were conducted in maths lessons of the "more-fixed mindset teachers" Ms Evans and Ms Walker (School 1 – Meadows School) as well as the "less-fixed mindset teachers" Ms Barns and Ms Green (School 7 – Woodpark School). Please note that teachers and schools are referred to with pseudonyms throughout this thesis. Relevant characteristics and means of the motivational beliefs of these four teachers are presented in Table 3.4. Please refer to Section 3.4.1.1 to see how these constructs were measured. The selected teachers as well as groups of 4 girls and 4 boys, 5 girls and 3 boys, 4 girls and 4 boys, as well as 2 girls and 2 boys from their classes, respectively, participated in interviews. Please note that it seemed to have been a coincidence that the selected more-fixed mindset teachers taught at the same school and that the selected less-fixed mindset teachers taught at the same school. Mindset scores were relatively similar across schools. In addition, I initially selected a more-fixed mindset teacher from a different school, but he did not want to participate in the observations.

Table 3.4

Characteristics of the Mixed-Methods Teachers

Variable	**Ms Evans**	**Ms Walker**	**Ms Barns**	**Ms Green**
Mindset compared average	More fixed	More fixed	Less fixed	Less fixed
School	1	1	7	7
Class taught	Top (1/5)	Bottom (4/5)	Middle (3/6)	Middle (4/6)
Education	Master's	PGCE	PGCE	Master's
Years of teaching maths	4	3	11	11
Additional subjects	-	-	-	Language
Fixed mindset	1	1.75	0	0.25
Mastery goal	3.33	3.67	3.67	3.00
Performance-approach goal	2.50	1.75	2.25	1.75
Performance-avoidance goal	3.67	4.00	3.33	3.00
Failure-to-ability attribution	2	1	0	2
Perceived ability/expectation	2.50	2.50	2.84	2.30

Note. Motivational beliefs were measured on a scale from 0 to 4.

3.4 Quantitative research instruments

Questionnaires were used to assess teachers' and students' self-reported motivational beliefs as well as teacher-reported and student-perceived FAPs. All questionnaires were piloted prior to the main study (see Section 3.7.1). Table 3.5 provides a short overview of the instruments that were administered in the main study.

Table 3.5

Overview of the Teacher and Student Questionnaires

Collection phase	Teacher questionnaire	Student questionnaire
Time 1 Motivational beliefs	• Mindset for students' ability • Goal orientation for teaching • Attribution for students' achievement • Perceived student ability/expectation	• Mindset • Achievement goal orientation • Achievement attribution • Self-concept
Time 2 Fixed-ability practice	• Tasks ○ Unsophisticated task differentiation ○ Unidimensional tasks • Grouping ○ Ability grouping ○ Stability • Evaluation ○ Performance over process recognition ○ Social comparison • Little autonomy ○ Little choice ○ Little promotion of self-regulation • Differential value • Little promotion of risk-taking ○ Little encouragement of help-seeking ○ Unconstructive handling of mistakes	
Time 3 Motivational beliefs	-	• Mindset • Achievement goal orientation • Achievement attribution • Self-concept

3.4.1 Motivational beliefs questionnaires

The motivational beliefs scales were mainly based on existing scales. To increase the credibility of the quantitative data, stringent criteria were applied to select appropriate and psychometrically sound instruments. Specifically, instruments were selected based on their validity and reliability (primarily internal consistency) in prior studies. Some items of selected scales were adapted to the needs of this study, and some response scales were modified to use a consistent scale throughout. All items as well as an indication of the adjustments made to the original scales can be found in Appendix A for the teacher (Table A.1) and student motivational beliefs questionnaires (Table A.2). Since modifying items and scales could compromise their reliability, a pilot study was conducted to assure that the instruments had adequate reliabilities (see Section 3.7.1).

An overview of all instruments comprising the motivational beliefs questionnaires is provided in Table 3.6 for teachers and Table 3.7 for students. This includes example items and information regarding the scales' validity and reliability in previous studies. Note that all teacher responses were given on a 5-point Likert-type scale from 0 = *strongly disagree* to 4 = *strongly agree* and all student responses were given on a scale from 0 = *not at all true* to 4 = *very true*, except for the perceived ability/expectation and self-concept scales, which is explained below.

3.4.1.1 Teacher motivational beliefs questionnaire

3.4.1.1.1 Teacher mindset

Teachers' beliefs about the malleability of their students' ability were assessed with an adapted version of the Implicit Theories of Intelligence Scale by Dweck (1999). The instrument consists of a fixed mindset subscale and a growth mindset subscale, which are comprised of four items each. The items were adapted in two ways. First, teachers were asked to make a judgement about their students' malleability of ability rather than their own, by replacing "you" with "my pupils." Second, to assess maths-specific mindsets, "intelligence" was replaced by "ability at maths." This is together with "[subject] ability" a common phrasing for subject-specific mindset items (e.g., Chen & Tutwiler, 2017; Dai & Cromley, 2014; Luo et al., 2014; Muenks et al., 2015). Please note that even though both scales were administered, only the fixed mindset items were used in analyses. This was done to avoid a social desirability bias caused by growth mindset items being highly compelling (see Lüftenegger & Chen, 2017; Pomerantz & Kempner, 2013).

Table 3.6

Example Items and Prior Psychometric Properties of Teacher Motivational Beliefs Questionnaire

Scale	Subscale (N items); example item	Validity and reliability in prior studies
Mindset[5]	**Fixed mindset** (4); "My pupils have a certain amount of ability at maths, and they can't really do much to change it."	• the original scale demonstrated construct validity (Blackwell et al., 2007; Dweck et al., 1995) • predictive validity (Blackwell et al., 2007; Good et al., 2003) • high internal reliability with α>.86 (Bråten & Strømsø, 2006) and α=.80 (Diseth et al., 2014) • acceptable test-retest reliability at two weeks (α=.77; Blackwell et al., 2007)
Goal orientation for teaching	**Mastery** (3); "It is important for me to continue to learn more about teaching maths." **Performance-approach** (4); "It is important for me to teach maths better than other teachers." **Performance-avoidance** (3); "It is important that my teaching abilities in maths are not inferior to that of most of my colleagues."	• the original scales demonstrated construct validity and adequate internal reliability with α=.76, α=.82, and α=.71 for the three subscales (Butler, 2007), which was confirmed in later studies (e.g., Retelsdorf et al., 2010) • revised scales also demonstrated construct validity and adequate internal reliability with α=.82, α=.81, and α=.70 for the three subscales (Shim et al., 2013)
Achievement attribution[6]	**Failure-to-ability attribution** (1); "If a pupil does poorly at maths, it is probably because he/she lacks ability at maths."	NA
Perceived ability/ expectation	(2); "How well do you expect [Pupil] to perform this year in maths?"	• demonstrated predictive validity (Bohlmann & Weinstein, 2013; Wigfield et al., 1997) • reliability measures not appropriate given that the scale consists of two items

[5] Fixed and growth mindset items were administered but only fixed mindset items were used in analyses as discussed in Section 3.4.1.1.1.
[6] Four attribution items were administered but only the failure-to-ability item was used in analyses as discussed in Section 3.4.1.1.3.

Table 3.7

Example Items and Prior Psychometric Properties of Student Motivational Beliefs Questionnaire

Scale	Subscale (N items); example item	Validity and reliability in prior studies
Mindset[7]	**Fixed mindset** (4); "I don't think I can do much to increase my ability at maths."	• demonstrated construct validity, predictive validity, and good internal consistency (α >.90, Castelle & Byrne, 2015)
Achievement goal orientation	**Mastery goal** (5); "It's important to me that I learn a lot of new maths concepts this year." **Performance-approach goal** (5); "It's important to me that other pupils in my class think I am good at my maths work." **Performance-avoidance goal** (5); "It's important to me that I don't look stupid in my maths lesson."	• The original validation study reported adequate internal reliability with α=.85, α =.89, and α=.74 for the three subscales (Midgley et al., 2000) • construct validity and adequate internal reliabilities were confirmed in later studies (e.g., Hackel et al., 2016; Yu & McLellan, 2019)
Achievement attribution[8]	**Failure-to-ability attribution** (1); "If I do poorly in maths, it's probably because I don't have a lot of ability at maths."	NA
Self-concept	(5); "Compared to other pupils in your class, how well do you expect to do in maths this year?"	• demonstrated construct and predictive validity (for a review, see Eccles et al., 2005) • adequate internal reliability with Cronbach's α ranging between .74 and .92 (Eccles et al., 1993; Eccles et al., 2005; Wigfield et al., 1997)

[7] Fixed and growth mindset items were administered but only fixed mindset items were used in analyses as discussed in Section 3.4.1.2.1.
[8] Four attribution items were administered but only the failure-to-ability item was used in analyses as discussed in Section 3.4.1.2.3.

3.4.1.1.2 Teacher goal orientation

Teachers' goal orientations were assessed using the Achievement Goal Orientations for Teaching scale. This instrument was originally developed by Butler (2007) and later revised by Shim et al. (2013). The revised version was used because it addresses teaching goals more explicitly, while Butler's scale focuses on reasons for feeling successful; thus, the former has higher face validity. It has three subscales reflecting the trichotomous model of goal orientations. The mastery goal subscale and the performance-avoidance subscales are comprised of three items each, and the performance-approach goal subscale is comprised of four items. The items were adapted to address maths teaching rather than teaching in general.

3.4.1.1.3 Teacher achievement attribution

Four items were administered to assess teachers' level of agreement to attributing good and poor maths performance to effort and ability, which were based on an instrument for kindergarten teachers' attributions for children's performance (Upadyaya et al., 2012). After further consideration, I decided to only use the failure-to-ability attribution item for analyses. Reasons include that this form of attribution is considered to be particularly maladaptive because it is a form of helpless attribution, and that this type of attribution has previously been linked to a fixed mindset (Blackwell et al., 2007; Hong et al., 1999).

3.4.1.1.4 Teacher perceived ability/expectations

Two questions were used to assess teachers' perceived ability and performance expectations for their students. Teachers were asked to rate each participating student in their class relative to their classmates on a 5-point scale from 0 = *One of the lowest performing* to 4 = *One of the highest performing*. This correspondents to the item used in the student questionnaire, which was taken from Eccles et al. (2005). The second item was "How well do you expect [Pupil] to do this year in maths?" This question was adapted from Wigfield et al. (1997). Responses were given on a 5-point Likert-type scale ranging from 0 = *Not at all well* to 4 = *Very well.*

3.4.1.2 Student motivational beliefs questionnaire

3.4.1.2.1 Student mindset

Student mindset was assessed with the Implicit Theories of Intelligence Self-Theory scale by Castella and Byrne (2015). The self-theory scale was chosen because it was found to predict students' goal orientation and performance attributions above and beyond Dweck's (1999) measure (Castella & Byrne, 2015). This instrument includes a fixed and a growth mindset subscale that are comprised of four items each. To assess maths-specific mindsets, "intelligence" was replaced by "ability at maths." Moreover, some items were simplified to be more appropriate for the age group. Although both scales were administered, only the fixed mindset items were used in analyses for reasons outlined above.

3.4.1.2.2 Student goal orientation

Students' goal orientations were assessed using the Personal Achievement Goal Orientations scale from the Patterns of Adaptive Learning Scales (PALS; Midgley et al., 2000). This instrument is one of the two most widely used measures for assessing goal orientation (Hulleman, 2010). I chose it over the other widely used measure, the Achievement Goals Questionnaire (AGQ; Elliot & Murayama, 2008), because it uses the trichotomous model of goal orientation and because it focuses on the more maladaptive appearance goals rather than normative goals. The instrument consists of 14 items that comprise three subscales. Mastery goals and performance-approach goals are assessed by five items each, and performance-avoidance goals are assessed by four items. Following the suggestion by Midgley et al. (2000), the items were adapted to be domain-specific, in this case, maths-specific. Additionally, a fifth item was added to the performance-avoidance subscale because its reliability was insufficient in the pilot study. However, this item was not used in the main analyses as the original scale reached sufficient reliability in the main study.

3.4.1.2.3 Student achievement attribution

The four items assessing students' attributions for their good and poor maths performance were developed to closely resemble the items used for teachers. Thus, they were also based on Upadyaya et al. (2012). Students' were asked to indicate how true it was for them that their good or poor performance are determined by ability or effort. As explained in the teacher section, only the failure-to-ability item was used in the analyses.

3.4.1.2.4 Student self-concept

To assess students' self-concept in maths, I used the subscale Ability/Expectation from the Ability Self-Perceptions and Subjective Task-Value scale for adolescents (grades 5 through 12; Eccles et al., 2005). It is comprised of five items with a 5-point Likert-type response scale. The items assess how good the students think they are at maths in general and in comparison to other students in the class as well as how well they think they have done in the past and will do in the future.

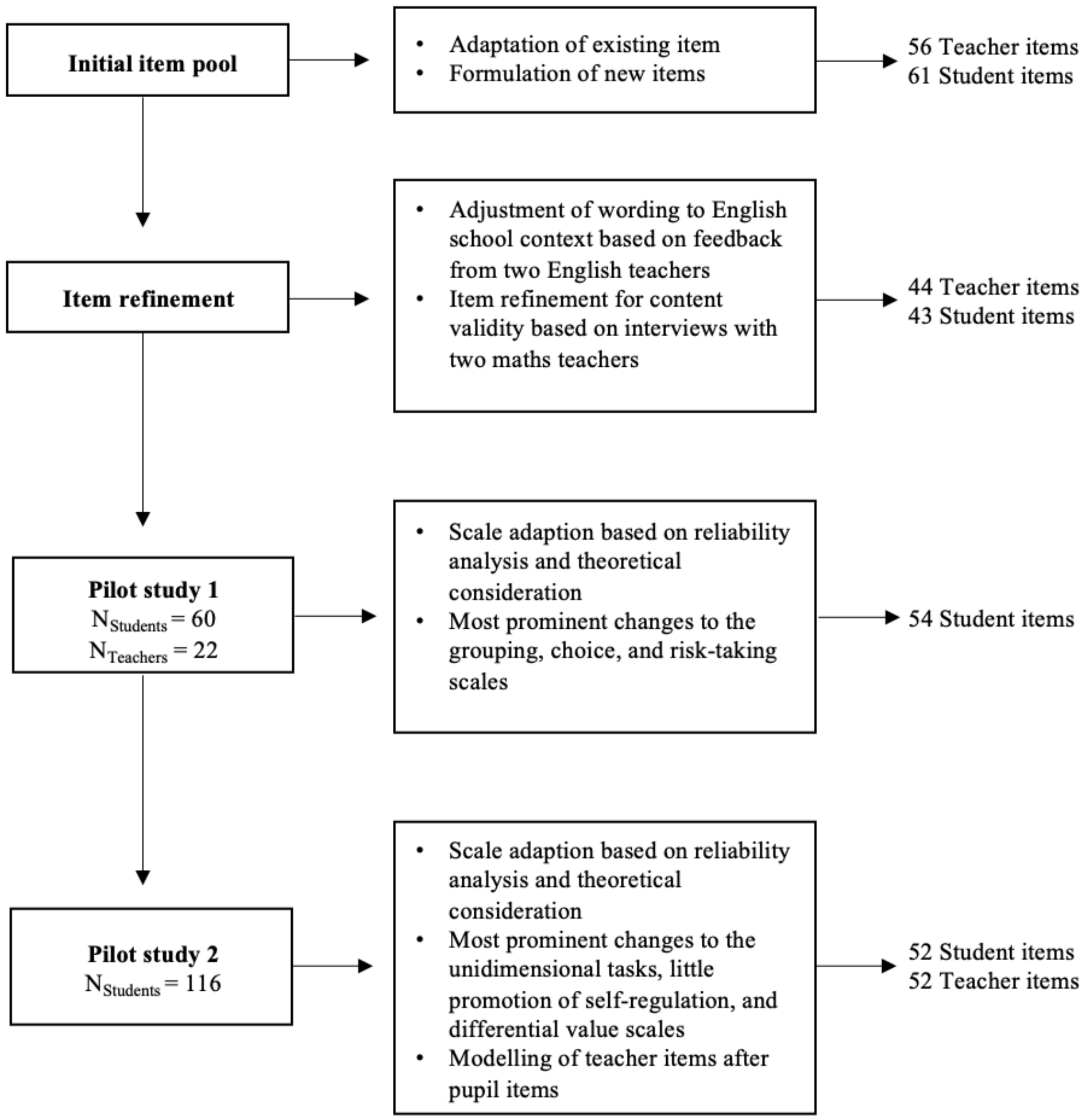
3.4.2 Fixed-ability practices questionnaire

As described in the introduction, no questionnaire for fixed-ability practices existed, which is why it was newly developed for this study. The development proceeded in several steps, and this process is graphically presented in Figure 3.2. The development of the FAPs questionnaire for teachers and students was guided by the characterisation of the six FAPs dimensions described in the literature review: tasks (unsophisticated task differentiation, unidimensional tasks), grouping (ability grouping, stability), evaluation (performance over process recognition, social comparison), little autonomy (little choice, little promotion of self-regulation), differential value, and little promotion of risk-taking (little encouragement of help-seeking, unconstructive handling of mistakes).

To develop the initial item pool, I first searched the literature for existing scales assessing these practices. When appropriate, items from these instruments were taken, adapted, and rearranged into the proposed categorisation. When no suitable items were found, new items were formulated based on the conceptual understanding of the respective dimensions. I initially took and adapted items from the TARGET questionnaires from Lüftenegger et al. (2017) and Vedder-Weiss and Fortus (2017). The FAPs questionnaires also included items from the Approaches to Instruction and the Perception of Classroom Goal Structures scales from the PALS (Midgley et al., 2000). In addition, the student scales included items from the Perceptions of Teacher Goals scale from the PALS. Items from the Perceived Teacher Support and Perceived Teacher Inhibition of Questioning and Help Seeking scale (Butler & Shibaz, 2008) were used to assess teachers' support of students' help-seeking. Notably, items were formulated so that they addressed concrete teaching practices (e.g., "When we make a mistake, my teacher stops others from making fun of it") rather than general classroom climate (e.g., "In our class, it's important that you don't make mistakes in front of everyone"; item taken from Midgley et al., 2000). This was done because this would enable providing more practical advice for teachers regarding the teaching practices they should and should not engage in.

Figure 3.2

Flowchart of the Development of the Fixed-Ability Practices Questionnaires



After developing the initial item pool, I further refined the items to ensure that they align with the experiences and language of teachers as well as that the scales have adequate content validity. Consulting members of the target population can provide important information about the ease with which items are understood and how relevant and representative they are of a given construct (Vogt et al., 2004). Hence, the initial items were first reviewed by two teachers to correct any language mistakes. I then conducted individual in-depth interviews with two additional teachers who taught maths at secondary schools. One teacher was female and had taught maths for approximately 21 years in the UK, and the second was male and had taught maths for four years in Canada. The interviews explored how teachers conceptualised the practice dimensions, what aspects they found representative, whether the initial items resonated with them, and how they would adapt the wording of existing items. The interviews also aimed to identify irrelevant or redundant items and to determine what additional items should be included. This was then followed by two pilot studies (see Section 3.7.1).

In the main study, 52 items were administered (see Appendix A, Table A.3). Corresponding to the FAPs conceptualisation, these were assumed to form six scales with multiple subscales each. Table 3.8 provides an overview of the administered scales, including the item numbers as well as example items for teachers and students. Teachers reported the frequency of their own practices and students reported the frequency of their maths teachers' practices on a 5-point Likert-type scale from 0 = *Never* to 4 = *Always*. Please note that only 32 items were retained for analyses based on the results of the factor analysis and that only student responses were used in the main analyses because the teacher scales did not reach adequate internal reliability (see Chapter 5). The final items used in the main analyses are presented in Chapter 5.

Table 3.8

Proposed Scales and Example Items of Fixed-Ability Practices Questionnaires for Teachers and Students

Proposed scale	Proposed subscales (N)	Teacher example item	Student example item
Tasks	Unsophisticated task differentiation (4)	"I give low achieving pupils easier maths tasks than others."	"In our maths lessons, all pupils get the same task (e.g., the same work sheet)."
	Unidimensional tasks (6)	"I assign maths tasks that have one correct solution."	"My maths teacher gives us tasks that need to be solved in one specific way."
Grouping	Ability grouping (4)	"When I arrange pupils into groups, I do it according to similar abilities at maths."	"When we do group work, my maths teacher puts pupils with similar ability at maths into one group."
	Stability (4)	"Over the course of the school year, I put the same pupils together into one group."	"During the school year, my maths teacher puts the same pupils together in one group."
Evaluation	Performance over process recognition (5)	"I praise pupils for putting a lot of effort into their maths work." (reversed)	"My maths teacher recognises us for trying hard." (reversed)
	Social comparison (4)	"I give special privileges or rewards to pupils who do the best maths work."	"My maths teacher tells us how we compare to other pupils."
Little autonomy	Little choice (5)	"I let my pupils choose with whom they want to work." (reversed)	"All pupils can choose between different questions when our maths teacher gives us tasks (e.g., tests and worksheets)." (reversed)
	Little promotion of self-regulation (5)	"I ask my pupils to assess their own maths work." (reversed)	"My maths teacher asks us to assess our own maths work." (reversed)
Differential value	Differential value (5)	"I carefully listen to the ideas and contributions of all my pupils." (reversed)	"My teacher seems to listen more carefully to pupils who do well in maths."
Little promotion of risk-taking	Little encouragement of help-seeking (4)	"I provide a lot of time for pupils to ask questions." (reversed)	"My maths teacher gives us a lot of time to ask questions." (reversed)
	Unconstructive handling of mistakes (6)	"I stress to my pupils that getting right solutions in maths is very important."	"My maths teacher tells us that mistakes are okay as long as we are learning from them." (reversed)

3.5 Qualitative research instruments

This section describes the interview schedules that were used for teachers and students. Please note that the classroom observations are described in the procedures (see Section 3.7.2.2.1) because observations were conducted to provide the video segments for the interviews rather than to be a data source that formed part of the analysis.

3.5.1 Interview schedule

Interview schedules were developed to guide me, the interviewer, during video-stimulated recall interviews that were conducted individually with teachers and with groups of students from the observed classes. Interviews with teachers aimed to provide greater insight into why teachers engage in FAPs. Interviews with students aimed to provide an understanding of how they perceive FAPs and, more specifically, what messages they draw from these for their motivational beliefs. In developing the interview questions, I was guided by Seidman's (2006) recommended interview techniques, including avoiding leading questions and using open-ended questions. Example interview schedules can be found in Appendix B for teachers (B.1) and students (B.2).

Please note that I conducted the interviews after I had observed several maths lessons, during which I developed good rapport with teachers and students. Accordingly, the tone of interviews was friendly and conversational, and the participants usually ate their lunch or snacks that I provided. I first always explained why I conducted the interview and checked whether teachers were happy to answer a lot of "why" questions to explain their engagement in different practices and whether students were happy to talk about how different teaching practices made them feel. Importantly, I always started teacher interviews with less sensitive practices or with situations in which the teacher seemed to engage in adaptive practices to set them at ease.

Interview schedules were individually adapted for each teacher and student group following observations. Hence, the order in which the teaching practices were addressed varied between interviews depending on what seemed most prominent during lessons. Situations in which practices relevant to FAPs occurred were identified and short video clips were selected from the lesson recordings. These were integrated into the schedules to explore why a teacher engages in that specific practice or what students think about it. For instance, for the dimension *evaluation*, I showed a video clip in which a teacher asked students to read out their test scores in front of the class. I then asked the teacher for her reasons for doing this.

Similarly, I asked her students how they felt about this. As another example, to discuss *unconstructive handling of mistakes*, I showed a video clip in which the teacher shakes her head when someone made a mistake and then moves on to another student. Again, I asked the teacher why she reacted this way, and I asked students what they thought about her reaction. When I did not observe any representative practice of a given dimension, I explored the dimension without video clips. Students were sometimes also asked to imagine their reactions to practices that did not occur in their classes. This was effective in pilot interviews because students readily discussed such practices, sometimes referring to other teachers that engaged in these practices. The aim was to identify potential explanations for relationships between FAPs and motivational beliefs in the overall sample, even if these did not occur in the mixed-methods strand. Aspects of interest that could not be captured on video, such as the bases for group compositions, were also discussed without showing video segments. For example, I asked teachers how they seated or grouped students, and I asked students whether they were paired with similar or different students. The analysis of the interview data and the applied coding scheme are described in Chapter 8.

3.6 Maths achievement

Key Stage 2 maths results as well as end of Year 7 maths grades were obtained from the school records. Since there was no standardised assessment at the end of Year 7, students' achievement from different schools could not be compared. Schools used different grading systems as well as different marking standards. To aid comparability, standardised z-scores for maths grades were used, whereby z-scores were computed for every school separately (Field, 2009). If different tests were used within the same schools, these were standardised separately as well given that different tests cannot be compared.

3.7 Procedure

3.7.1 Pilot studies

Two pilot studies were conducted to test whether the chosen research methods and instruments would produce the data necessary to answer the research questions as well as to reduce any procedural weaknesses prior to the main study. These are described in detail in Appendix C. In the first pilot study ($N_{Student}$ = 60; $N_{Teacher}$ = 22), I tested all research methods and instruments, including the student and teacher motivational beliefs and FAPs questionnaires, classroom observations, as well as student and teacher interviews. It was carried out at an all-girls secondary school. Reliabilities of the motivational beliefs scales were adequate for both teachers and students except for the performance-avoidance goal scale for students. Accordingly, an additional item was formulated and administered during the main study. Several FAPs scales had problematic internal reliabilities, which is why major revisions were performed (see Table C.5). Piloting the qualitative methods showed that these were appropriate for the purposes of the study, and only small adjustments were made to procedural aspects, such as using a smaller, less intrusive video camera.

After making initial revisions to the FAPs scales based on Pilot Study 1, a second pilot study ($N_{Student}$ = 116) was carried out to test this instrument. Due to the difficulty of obtaining a second teacher sample, only the student questionnaire was piloted for a second time. The analysis of the pilot data was then used to finalise the student FAPs questionnaire (see Table C.7), and the final student items were subsequently used to revise the teacher items.

Overall, the pilot studies demonstrated that the chosen methods could produce the data necessary to answer the research questions. Only minor adaptations were necessary for the motivational beliefs scales and qualitative data collection procedures, whereas the FAPs scales required more substantial revisions.

3.7.2 Procedures of the main study

The data were collected in three phases; one at the beginning, one in the middle, and one towards the end of the school year. At the beginning of the year, the head of maths sent opt-out letters to all parents or guardians, which needed to be returned if they did not wish their child to participate. Only students who were not opted-out were invited to participate. Students and teachers were asked to sign informed assent and consent forms, respectively, during the first session. All information letters and consent forms can be found in Appendix D. Students completed paper-based questionnaires during their maths lessons in their usual classrooms, while teachers completed all questionnaires online at their own time. The items were presented in a random order across the whole questionnaire so that items from the same scales were not all presented together. Only perceived ability/expectation and self-concept items were presented together at the end of the motivational beliefs questionnaires because of their specific response options. The order of items was the same for all participants. Please note that teachers sometimes read out the questions to some students or the entire class, which was mostly the case in bottom set classes.

3.7.2.1 Time 1: Teacher and student motivational beliefs questionnaires

At Time 1, the beginning of the school year, the motivational beliefs questionnaires were administered to students and teachers. I visited each class and introduced myself, and explained the study and procedure, confidentiality of data, and their right to decline participation or skip questions they did not want to answer. Students were encouraged to ask questions and were then invited to sign the informed assent form for the entire study, provided their parents did not opt them out. Students then completed the motivational beliefs questionnaire, comprised of the mindset, goal orientation, attribution, and self-concept scales. Students were allowed to ask questions when they did not comprehend an item. I was present during the whole administration whenever possible. However, when questionnaires had to be administered to multiple classes during the same lesson, I only gave instructions and took informed assent and then moved on to the next class. I provided envelopes for completed questionnaires to ensure that teachers could not read students' responses. All student questionnaires were administered during the first five weeks of the school year. The first session lasted approximately 20 minutes.

A few days after students completed the questionnaires, teachers received links to an online questionnaire assessing their mindset for their students' ability, goal orientation for teaching, achievement attribution, and perceived ability/expectations for their students. This delay was necessary because I used students' names provided on questionnaires to finalise the perceived ability items, which teachers completely individually for all of their participating students. Teachers gave informed written consent either during students' questionnaire administration or online. The completion time for the teacher motivational beliefs questionnaire varied. Some teachers finished in less than 5 minutes and others took up to 20 minutes, but most teachers took around 10 minutes.

3.7.2.2 Time 2: FAPs questionnaires, observations, and interviews

3.7.2.2.1 Procedures for the mixed-methods strand

Observations and interviews were conducted in the classes of the four selected teachers for the mixed-methods strand. As video-recordings and interviews are more sensitive than questionnaires, parents were provided with the opportunity to request that their children were not being recorded or did not participate in interviews (see Appendix D, D.2), and children could also decline. For each of the four teachers, I observed five to six mostly consecutive maths lessons over a two-week period. I positioned myself as well as the camera at the back of the classroom to cause as little intrusion as possible. Students who did not want to be recorded were seated at the other side of the classroom so that they were not captured by the camera. Teachers were also equipped with a voice recorder, which they carried around their neck, to record conversations with individual students. The footage obtained on the first day of observation was not used in the analyses as these recordings were used to familiarise teachers and students with the procedure. During observations, I took field notes, including information about the overall lesson structure and activities, teacher verbal comments, as well as indications of where FAPs occurred. A short example of these notes is provided in Appendix E.

Following observations, individual interviews were held with each of the four teachers. In addition, one boy- and one girl-group from each observed class were interviewed. Participants were reminded that individual answers would not be shared and that the recordings would only be accessed by me and my supervisor. All interviews were voice-recorded after permission was given by participants. Each interview started with warm-up questions to set interviewees at ease, such as asking them how they experienced the observations. In teacher interviews, this was followed by questions exploring what practices they used and why. In student interviews, this was followed by how these practices affected students' motivational beliefs. This involved showing video clips of situations in which FAPs occurred. It was common that some students were more eager to answer questions and share their opinions than others, but I tried to engage with all students. The interviews took place in empty classrooms during lunch breaks or enrichment time. Teacher interviews lasted between 38 and 59 minutes ($M = 46$) and student interviews lasted between 32 and 47 minutes ($M = 41$).

3.7.2.2.2 Procedures for the entire sample

All participants completed the FAPs questionnaire in the second phase. Students of the mixed-methods strand did so only after observations and interviews took place. On average, the FAPs questionnaire was administered approximately 6 months after the motivational beliefs questionnaire. The administration procedure did not differ from the first administration except that informed assent was not taken again given that assent had been granted for the entire study at Time 1. For two classes, I was not present during administration, and teachers read out my instructions. At one of the schools, the headteacher decided that the questionnaire administration should start at the beginning of the maths lesson. Hence, teachers read out my instructions and I moved between classes to answer questions. This second session for students took approximately 20 minutes. Online links to the FAPs questionnaires were sent to teachers after their students completed the questionnaires. Teachers completed the questionnaire in approximately 10 minutes.

3.7.2.3 Time 3: Student motivational beliefs questionnaire

At the third time point, towards the end of the school year, all students completed the same motivational beliefs questionnaire as at the beginning of the school year. The time gap between the two assessments of motivational beliefs was approximately nine months. In some classes of one school, I administered the second beliefs questionnaire a few weeks after the Easter break because they had a teacher change around that time, resulting in a time gap between motivational beliefs questionnaires of about seven months. The sessions lasted approximately 10 minutes. Due to practical reasons, I was not present during the questionnaire administration for some classes, which is when teachers administered the questionnaires using my instructions (see Appendix F).

3.8 Ethical considerations

The ethical guidelines set out by BERA (British Educational Research Association, 2011) and BPS (British Psychological Society, 2014) were considered during all stages of the study design and implementation. The study received approval by the Ethics Committee of the Faculty of Education. In addition, individual research guidelines of participating schools were followed. As described, an opt-out letter was sent to parents and guardians. Students were only invited to participate and provide informed assent if their parents did not opt them out. Informed consent was also obtained from all teachers. Information regarding the purpose and methods of the study were provided but the exact hypotheses were withheld to avoid influencing responses. Consent/assent forms emphasised that participation was voluntary and that participants could withdraw at any stage. Parents and participants were invited to ask the researcher questions and discuss any concerns. After the study, schools were fully debriefed and provided with a report of the results.

Data were anonymised by assigning all participants an identification number, and names and responses are being stored separately. The IDs link schools, teachers, and students. Participants' names were required to merge data from the different phases as well as to match questionnaire responses with achievement data. Names were removed after all data had been collected. Pseudonyms were used for participants and schools in the mixed-methods strand and no identifying contextual or geographical information was disclosed. Anonymity was a particular concern for the mixed-methods strand because videoing and interviews made anonymisation almost impossible (Wiles, 2013). Thus, the video material is accessible to only me and my supervisor, and students' faces will be blurred when video segments are shown to other academic audiences. Interviewees were asked not to share individual comments outside of the group. When disseminating this research, group rather than individual results will be presented whenever possible. Care is taken to store all data securely following the General Data Protection Regulation (2018), which was implemented after the data had been collected. Data are used for academic purposes only.

No physical harm was expected to result from participation, and psychological or emotional harm was thought to be minimal if not absent. Nevertheless, some participants might have experienced distress or discomfort (Merriam & Tisdell, 2015). To reduce emotional distress, participants were informed that they were not obliged to answer questions they did not want to answer. In addition, I tried to build good rapport with the classes to put participants at ease during observations and interviews. Lastly, to reduce the time investment of all participants, the overall length of the questionnaires and interviews was kept as short as possible without unnecessarily compromising the data.

Chapter 4
General Quantitative Analytical Strategy and Analysis Approach for FAPs Dimensionality (RQ 1)

This chapter first presents the overall analytical approach to address research questions (RQs) 1 to 3. It describes the data structure, use of multilevel modelling, and handling of missing data and careless responses (Section 4.1). It then moves on to a detailed explanation of how the quantitative data were analysed to answer RQ 1, that is, to determine the dimensionality of student-perceived fixed-ability practices (FAPs, Section 4.2). The results for RQ 1 are presented in the next chapter, and the analyses and results for the subsequent RQs are presented in later chapters.

4.1 The general quantitative analytical strategy

4.1.1 Overview of the general quantitative analytical strategy

First, the dimensionality of FAPs (RQ 1) was established based on student data using exploratory first-order structural equation modelling (ESEM), confirmatory ESEM, hierarchical ESEM as well as bifactor ESEM, which are all described in this chapter. Teacher data was not used due to the insufficient sample size. Second, I initially planned to examine the relationships between teachers' motivational beliefs and their engagement in student-perceived and teacher-reported FAPs (RQ 2) using regression analyses. However, regression analyses were not warranted because of a lack of significant correlations and a relatively small teacher sample size. Third, to analyse longitudinal relationships between student-perceived and teacher-reported FAPs and students' motivational beliefs and achievement (RQ 3), I first examined longitudinal measurement invariance for motivational beliefs and then estimated multivariate cross-classified multiple membership random effects models (see Chapter 6 and 7 for analyses and results, respectively).

Statistical analyses were performed in three different software packages. Descriptive statistics were analysed in SPSS (IBM Corp., 2017) and R (R Core Team, 2017). I analysed the dimensionality of FAPs in Mplus 8.6 (Muthén & Muthén, 1998-2017). Lastly, the relationships between FAPs and students' motivational beliefs and achievement were examined in R using the brms package (Bürkner, 2017, 2018), which is a Stan interface[9] (Carpenter et al., 2017).

4.1.2 Data structure and multilevel modelling

The student data in this study was hierarchical because students (Level 1) were nested within teachers (Level 2) and classrooms (Level 2). Accordingly, the data were analysed using multilevel modelling to take clustering into account rather than to assume that observations and errors are independent, which are assumptions of single-level analysis techniques (Tabachnik, 2014). I did not include the school level in the multilevel models because the data were gathered at only seven secondary schools, which was an insufficient size for forming a hierarchical level since only a very rough estimate of the between-school variance could have been obtained (Van den Noortgate et al., 2005).

In traditional multilevel modelling, units at a lower level (e.g., students) belong to only one element of a higher-order unit (e.g., classrooms). This is referred to as a purely hierarchical structure. In this study, the structure was not purely hierarchical because some classes, and thus students, were taught by two or even more maths teachers throughout the year, some classes changed teachers, and some students changed classes and thus also teachers (see Figure 4.1 for a schematic representation of the data structure). When lower-level units belong to two or more higher-order units that have no pure hierarchical structure (here: some classes were taught by multiple teachers and some teachers taught multiple classes as opposed to one class was taught by only one teacher and one teacher taught only one class), the data are cross-classified (Beretvas, 2014; Browne et al., 2001; Goldstein, 2003). In many studies, the classroom and teacher levels are completely confounded so that it is impossible to distinguish the two. However, because some teachers taught multiple classes in this study, it became possible to account for both levels, although the distinction is not perfect as many teachers taught only one class.

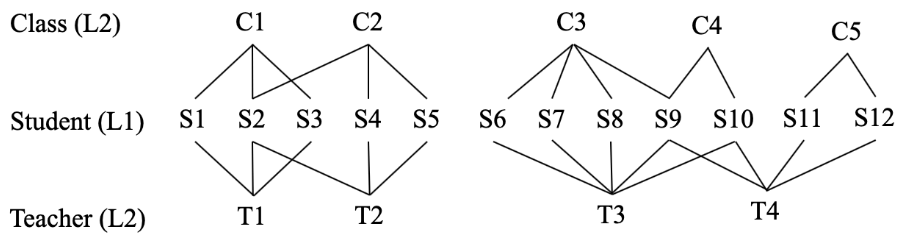
In addition to the cross-classification, the data also have a multiple membership structure because some students changed their classes and thus belonged to multiple classrooms and teachers. In past research, this complex structure has often been ignored by only considering the classroom or school that a student belonged to for most of the time or excluding students that belonged to multiple groups. However, this has severe shortcomings as failing to account

[9] All computations are conducted in Stan (Stan Development Team, 2018a) while the syntax is written in R language. Hence, the package requires a C++ compiler. Stan is a platform for statistical modelling and high-performance statistical computation, including full Bayesian statistical inference.

for this structure can result in negatively biased standard error estimates, which will then inflate the Type I error rate of associated statistical tests (Browne et al., 2001; Fielding & Goldstein, 2006; Meyers & Beretvas, 2006; Raudenbush & Bryk, 2002), as well as inaccurate variance component estimation (Fielding & Goldstein, 2006; Rasbash & Browne, 2008). Accordingly, I decided to take the complex structure into account to estimate more accurate models to answer RQ 3 (see Chapter 6). Note that this was not necessary for analysing the factor structure of FAPs because students reported practices for only one teacher. Hence, a traditional two-level model was used (see Section 4.2).

Figure 4.1

Network Graph of the Cross-Classified Multiple Membership Structure



4.1.3 Missing data and careless responses

Missing data were analysed separately for the three time points. Only students who participated at a given time point were considered in the missing data analysis for that particular time. This was done because participating at only some of the time points could mostly be explained by students' absence during data collection rather than a refusal to participate. Accordingly, students who completed only some questionnaires were unlikely to differ systematically from those who completed all. Students who gave no or invalid responses (e.g., choosing more than one response option) to more than 25% of the items of a given questionnaire were excluded from analyses because this indicates problems, such as a lack of motivation or having difficulty understanding the items. This was the case for eight participants at Time 1 and Time 2, and six participants at Time 3. Notably, missing responses for the grouping items at Time 2 were not counted because students were instructed not to respond to these if they did not engage in group or pair work.

Before imputing missing data, additional participants were excluded because of careless responding. Careless responding was examined using the longstring method, that is, the examination of multiple consecutive identical responses, also called straightline responses (DeSimone et al., 2018). No established cutoff criteria exist, but it is recommended to inspect frequency distributions of longstrings to decide on critical values (DeSimone et al., 2018). For all three time points, the longstring frequency drastically dropped going from five to six identical consecutive responses, which is why six consecutive responses could be considered as unusual. However, due to the relatively low percentage of participants showing this pattern (Time 1: 6.90%, Time 2: 5.90%, Time 3: 7.80%), their data were not expected to influence the results even if they were truly straightline responses. DeSimone et al. (2018) showed that partial straightlining (around 20% of data are straightlining responses) did not meaningfully influence responses if this was the case for less than 10% of cases. Hence, I decided not to exclude participants based on that cutoff, but I inspected responses from all participants who gave such identical consecutive answers in more detail. Most of these seemed genuine because there were few contradictory responses except for sometimes agreeing to both growth and fixed mindset items. Yet, one participant each was excluded at Time 1 and Time 2, and three participants were excluded at Time 3 because of highly suspicious longstrings, including giving the same response to all but the first few or last few items in a questionnaire.

After excluding the participants as described above, the overall percentage of missing data was 1.41%, 1.08% and 0.73% for Time 1, Time 2, and Time 3, respectively. The percentage of missing data for individual items ranged from 0.1% to 4.0%, from 0.1% to 3.7%, and from 0.1% to 1.4% for the three time points, respectively. At this stage, I decided to exclude all grouping items from the FAPs questionnaire from analyses. These items were not only problematic because many classes did not engage in group or even pair work but also because many students from classes that did not engage in such work responded to the items regardless. Hence, the missing data analysis and all subsequent analyses excluded these items.

Whether missing data was completely random or not was tested using Little’s MCAR test in SPSS. Only missingness at Time 1 was completely random (χ^2 = 2830.764 (2739), p = .108) but not at Time 2 (χ^2 = 5736.460 (5103), p < .001) or Time 3 (χ^2 = 2476.604 (1866), p < .001). It was thus important to establish whether missingness was not at random (MNAR) or at random (MAR). Accordingly, I examined whether missingness at certain variables could be explained by other variables in the dataset. I expected that the presence of multiple items assessing the same construct as the missing item would be able to explain missingness. To test this, participants were divided into those with and without missing responses for each variable, and t-tests were performed to test whether missingness could be explained by other variables in the data set. This was indeed found for most of the variables, which supports MAR. While MAR cannot be proven with complete certainty, these results as well as the fact that missingness overall was very small in this study seemed sufficient for using imputation techniques. Schafer (1999) argued that a missing rate of 5% or less is inconsequential. Likewise, Bennett (2001) stated that statistical analyses are only likely to be biased when more than 10% of data are missing.

To impute missing values, I used Bayesian single imputation, which takes all available information from the observed data into account and is thus similar to Full Information Maximum Likelihood (FIML) estimation. Other missing data approaches, including FIML and Multiple Imputation (MI), have been considered but single imputation using Bayes was found to be the best approach given the statistical analyses that were subsequently performed. FIML was not optimal because important fit indices are not being provided for factor analysis when used with ordinal data. Using single imputation with Bayes enabled me to use the Weighted Least Squares Estimator (WLSMV) in Mplus, which treats data as ordinal. This estimator uses pair-wise present, leading to a decrease in sample size, which is why it was important to impute data beforehand. Another disadvantage of FIML is that only those variables that are included in a given model are being used for estimation. In contrast, with imputation, I could include all relevant items, which means I could use a greater set of variables to estimate missing values (see Asparouhov & Muthen, 2010). MI would have been ideal because it reflects the uncertainty of imputed values, but this could not be used because some of the planned analyses could not yet deal with multiple data sets. Because the proportion of missingness was small, underestimation of standard errors was unlikely to be a concern even when using single imputation (Sainani, 2015).

Missing data imputation was performed in Mplus using Bayesian imputation based on the unrestricted variance covariance model and taking the clustering structure into account. Missing values were imputed on the item- rather than scale-level because this was found to drastically improve precision (Gottschall et al., 2012). All items were specified as categorical (ordinal). Imputation was conducted for every time point separately, whereby all items administered at the given time point were used in the imputation. Additionally, sex and Key Stage 2 maths performance were included in all imputation models but were not imputed themselves.

4.2 Analyses for research question 1

This section describes the analyses relating to RQ 1: What are fixed-ability practices? More specifically, what are the individual dimensions of fixed-ability practices, at least as perceived by students, and is there an overarching fixed-ability dimension?

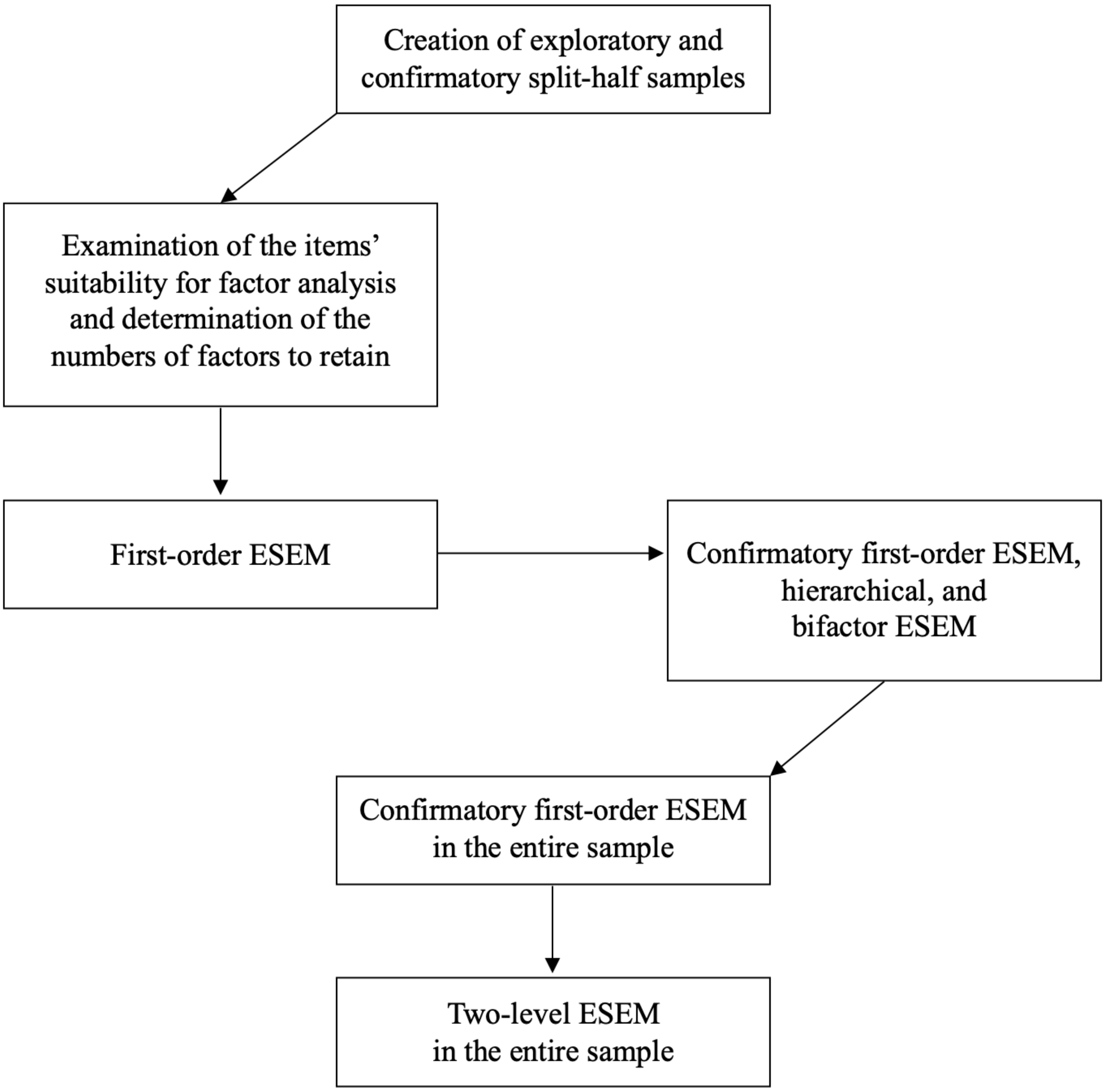
4.2.1 Analyses overview

Determining the factor structure of FAPs was based on the student data alone because the teacher sample was too small for factor analysis. The factor analysis proceeded in several steps, which are visually presented in Figure 4.2. First, the sample was randomly split into two halves to form an exploratory (n= 411) and a confirmatory (n = 420) sample. This is a common approach to exploring and validating a factor structure using only one data set (Hair et al., 2014). To aid analyses, nine students were moved between samples after the random allocation, so that a minimum of five students belonged to each cluster[10]. The first split-half sample was used to explore the first-order factor structure of FAPs at the student level using exploratory structural equation modelling (ESEM; Asparouhov & Muthén, 2009). Here, ESEM was used in its most basic form, which is identical to the traditional exploratory factor analysis (EFA; Marsh et al., 2014). In contrast to EFA, ESEM can also be used in a confirmatory approach, which was used with the second split-half sample to confirm the student-level first-order factor structure. Confirmatory ESEM rather than traditional confirmatory factor analysis (CFA) was conducted because ESEM tends to represent the true factor structure more adequately by allowing for cross-loadings and deflating the factor correlations (e.g., Asparouhov & Muthén, 2009; Marsh et al., 2014; Morin et al., 2016). The approach is described in more detail below. To examine whether the FAPs dimensions have a hierarchical or general factor rather than only a simple first-order structure, I also estimated a hierarchical and bifactor model and compared the three models. Based on the best fitting model at the student level, the classroom- or between-level factor structure was explored using the whole sample. Descriptive statistics and psychometric properties of the student FAPs scales were then examined using the final model. Lastly, the reliabilities of the corresponding FAPs scales using the teacher data were examined.

[10] A given teacher in a given classroom formed one cluster.

Figure 4.2

Steps in Examining the Dimensionality of Student-Perceived Fixed-Ability Practices



Individual items were treated as ordered-categorical variables utilising the Means and Variance Adjusted Weighted Least Squares Estimator (WLSMV) because items had a five-point Likert-type response scale. While it is quite common to treat Likert-type data with five or more response options as continuous, this was not appropriate because FAPs items were strongly skewed with many responses falling on the extreme ends of the scale. Treating such data as continuous using maximum likelihood approaches can lead to bias in estimates and standard errors as well as misleading fit indices (Muthén & Kaplan, 1985). This being said, the WLSMV estimator assumes that the underlying latent constructs are continuous. Later analyses using composite scores for each construct treated the data as continuous because combining Likert-type items into a scale leads to interval data (for a discussion, see Carifio & Perla, 2008).

4.2.2 Model evaluation criteria

I evaluated model fit using several different indices, including the χ^2 Test of Model Fit, the Tucker-Lewis Index (TLI), Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA) with 90% confidence interval, and Standardised Root Mean Square Residual (SRMR). Following common interpretation guidelines (e.g., Browne & Cudeck, 1992; Hu & Bentler, 1999; Marsh et al., 2005; Marsh et al., 2004), values for the CFI and TLI greater than .90 and .95 are thought to indicate adequate and excellent model fit, respectively. For the RMSEA, values smaller than .08 or .06 indicate acceptable and excellent fit, respectively. SRMR values smaller than .08 indicate good fit. Note that the Model Chi-Square results need to be considered with caution as they are sensitive to sample size and normality of the data; violations of normality can result in model rejection despite good fit (McIntosh, 2007).

4.2.3 Exploratory factor analysis in the exploratory sample

As mentioned above, all items referring to grouping practices were excluded from the analyses. I then examined whether any of the remaining 44 items were unsuitable for factor analysis because they were not sufficiently correlated with other items belonging to the same conceptualised dimension. I used a loose cutoff of $\rho = .30$, because correlations this small indicate that the items do not measure the same construct and are thus often excluded from factor analysis (Hair et al., 2014; Jacobs et al., 2017). Nevertheless, rules of thumb should always be used with caution, which is why I also used personal judgment when making decisions about the exclusion of items, as will be described. Following the examination of correlations, I tested the factorability of the remaining items in SPSS using the Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test of sphericity. This was followed by examining the optimal number of extracted factors using a scree plot obtained in SPSS and parallel analysis[11] conducted in R.

ESEM was performed to first establish the student-level first-order factor structure of the FAPs items with the exploratory split-half sample. Clustering was taken into account using the complex structure option, which adjusts standard errors and the chi-square statistic.

---

[11] Parallel analysis compares eigenvalues extracted from the actual data and eigenvalues from a large number of random correlation matrices generated with Monte Carlo simulation. The number of factors is retained where the actual eigenvalue is higher than that of the simulated data sets (Horn, 1965).

Oblimin rotation of the factor loading matrix was used to obtain an oblique factor solution, that is, one in which factors are allowed to correlate. Items were removed iteratively based on cross-loadings (loading higher than .32 on two or more items, Costello & Osborne, 2005). Rather than basing the final factor solution on loadings alone, conceptual considerations were also taken into account because factors have to be interpretable; that is, the content of each retained item must be representative of the construct they are supposed to measure (Matsunaga, 2010; Raubenheimer, 2004). Thus, all items comprising one factor were inspected to identify the common underlying content, and items that did not belong were removed. In doing so, the initial conceptualisation of FAPs was also considered.

4.2.4 Confirmatory ESEM in the confirmatory sample

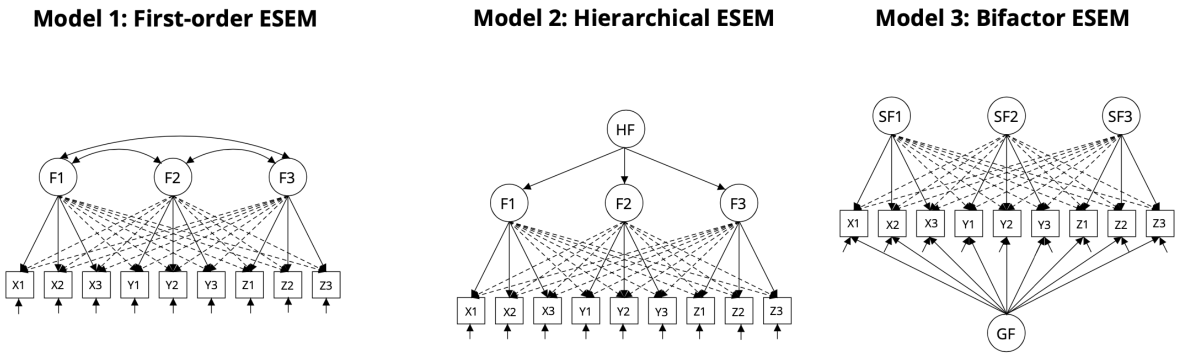
The student-level first-order factor structure obtained in the exploratory analysis was then tested using confirmatory ESEM with the second split-half sample. Confirmatory ESEM was conducted using target rotation to specify the priori hypothesis regarding the factor structure (Asparouhov & Muthén, 2009), which is explained in more detail below. In addition, a hierarchical factor model and a bifactor model were estimated. The three models were then compared to examine whether there was a hierarchical or general factor underlying all items. A simplified graphical representation of all three models is shown in Figure 4.3. In a classical hierarchical factor analysis, each item is specified as loading on its specific subscale (i.e., the first-order factor), and each first-order factor is specified as loading on a higher-order factor (see Morin et al., 2016). The relationship between an item and the higher-order factor (HF) is represented as mediated by the first-order factor (F). In contrast, in a bifactor model, items load directly on both a general factor (GF) and the first-order factor, called a specific facet (SF; see Morin et al., 2016). It thus tests if a general factor exists as a unitary dimension that underlies the responses to all items and coexists with multiple specific facets. The specific facets are defined by the part of the items that remains unexplained by the general factor.

4.2.4.1 First-order confirmatory ESEM

A first-order ESEM model using target rotation was estimated to test the student-level factor structure obtained in the exploratory analysis. The expected patterns of associations between items and factors were specified based on the results of the exploratory ESEM. More specifically, loadings for items belonging to a given factor were freely estimated while all other item loadings were targeted to be close to zero. The factor structure was then further refined based on factor loadings.

Figure 4.3

Simplified Graphical Representation of the Exploratory Structural Equation Models



Note. ESEM = Exploratory structural equation modelling

4.2.4.2 Hierarchical ESEM

To test whether there was a higher-order FAPs factor, a hierarchical ESEM model was estimated using ESEM-within-CFA as shown by Morin et al. (2016)[12]. The final first-order ESEM model was re-expressed using CFA, for which no rotation is necessary. This means that the exact values of the non-standardised loadings and cross-loadings from the first-order ESEM model were used as start values. For identification purposes, the variance of the higher-order factor was constrained to be one, whilst the other factor variances were freed. In addition, five loadings for every factor were fixed to be exactly the value that was estimated in the first-order ESEM, specifically, two large target loadings and three small non-target loadings. Lastly, the first-order factors were specified to define the higher-order factor.

4.2.4.3 Bifactor ESEM

A bifactor ESEM model was estimated to test whether there was a general FAPs factor besides the specific facets, using orthogonal bifactor target rotation. The specific facets were defined by main loadings from their respective items, while all cross-loadings were estimated but targeted to be as close to zero as possible. The general factor was defined through main loadings from all items and was part of the same set of ESEM factors as the specific facets (i.e., non-hierarchical).

[12] Some adaptations have been made based on recommendations by the Mplus support team, in particular Linda Muthén, during a private communication.

4.2.4.4 Model comparison

To compare models, I examined differences between fit indices. Models can be said to fit similarly well when differences in the CFI are < 0.01 and increases in RMSEA are < 0.015 (Chen, 2007; Cheung & Rensvold, 2002). Changes in the TLI of < 0.01 suggest a similar fit for models having a complex structure (Marsh et al., 2009; Morin et al., 2013).

4.2.5 Two-level factor analysis in the entire sample

Even though students reported their individual perceptions of their teachers' practices, it was expected that there was also agreement amongst students reporting practices of the same teacher in a given class. Thus, practices are not only an individual-level construct, but responses can (or even should) be aggregated to the classroom level, reflecting the joint perceptions of students (Lüdtke et al., 2009). Importantly, the structure at the student (within) and classroom (between) levels are not necessarily the same but often differ, whereby it is common that there are fewer dimensions at the classroom level (Lüdtke et al., 2009; Steuer et al., 2013). Before using the aggregated scores in further analyses, it is necessary to test whether there is sufficient agreement between students who rate the same teacher in a given class (Lüdtke et al., 2009), which I did following factor analysis using intraclass correlation coefficients. To examine the structure of FAPs at the classroom level, I performed a two-level factor analysis. This analysis was conducted using the entire student sample. Combining split-half samples was important considering the sample size at the classroom level.

As a first step, the student-level first-order structure was confirmed again using confirmatory ESEM with target rotation. Subsequently, two-level ESEM with oblimin rotation was performed to assess the number of factors at the classroom level. I estimated a series of models with the exact number of factors as in the final student-level structure at the student level and varying numbers of factors at the classroom level. I also estimated a model unrestricted at the classroom level to obtain a reference for the fit of the previous models.

Chapter 5
Results for FAPs Dimensionality (RQ 1)

This chapter presents the results of the analyses described in the previous chapter to answer the first research question (RQ), namely, "What are fixed-ability practices? More specifically, what are the individual dimensions of fixed-ability practices, at least as perceived by students, and is there an overarching fixed-ability dimension?" I first present results of the exploratory (Section 5.1) and confirmatory factor analyses, including the hierarchical factor model and bifactor model (Section 5.2), to determine the dimensionality of fixed-ability practices (FAPs) at the student level. Subsequently, results for the two-level factor analysis are presented to determine the number of factors at the classroom level (Section 5.3). The psychometric properties of the final FAPs scales for students (Section 5.4) and teachers (Section 5.5) are then described. Finally, a brief summary on the results for RQ 1 is provided (Section 5.6).

5.1 Results of ESEM

All items and their identifiers are presented in Appendix A (see Table A.3). Examination of the Spearman's Rho correlation matrix using observed data showed that the *unidimensional tasks* items did not correlate sufficiently high to be considered for factor analysis. Items TU2[13] and TU6 did not correlate above .30 with any other item; items TU4 and TU5 were only correlated with each other (ρ = .32); items TU1, TU3, TU4, and TU5 were correlated above .30 with a few items from other dimensions but were excluded nonetheless because it was unclear how they would fit together conceptually. Thus, all items from the unidimensional tasks scale were removed. This was not entirely surprising because this scale has been completely revised after the second pilot study (see Appendix C), and there was not enough time to pilot it again before the main study. All items relating to *performance over process recognition* correlated above .30 with at least one other item of the same subdimension but the correlations were not particularly strong, which is why these might not form a stable factor in the final solution. Yet, the items were retained at this stage. All five items relating to *choice* were excluded from factor analysis. Only item AC4 correlated sufficiently high with other items, specifically with AC2 (ρ = .53) and AC5 (ρ = .33). This was surprising as most items were taken from an existing scale, and only small adaptions

[13] Item identifiers are a combination of abbreviating the overarching dimension and the subdimension followed by the specific item number. For example, TU2 stands for Tasks – Unidimensional tasks – 2.

were made to simplify the wording. Item ASR1, relating to *promotion of self-regulation,* was excluded because it did not correlate above .30 with any other remaining item. Similarly, item RH1, relating to *encouragement of help-seeking*, was deleted as it did not correlate above .30 with any other item.

Thirty-one items were retained for factor analysis. The Kaiser-Meyer-Olkin measure of sampling adequacy was .91, and Bartlett's test of sphericity was significant (χ^2 (465) = 4271.72, $p < .001$). This indicated that factor analysis could be performed because the proportion of variance that might be common variance was sufficiently high and because the correlation matrix was not an identity matrix. Exploratory structural equation modelling (ESEM) was performed to first establish the student-level first-order factor structure of the FAPs items with the exploratory split-half sample. Please note that items were entered into all factor analyses without being reverse scored so that not all dimensions reflected maladaptive practices yet. After the factor structure had been established, items were reverse scored were appropriate, so that all dimensions reflected maladaptive FAPs. The scree plot indicated that between four and five factors should be retained, and the parallel analysis suggested that four factors should be retained. Hence, initial solutions for four and five factors, as well as six factors, were examined. The five-factor solution was preferred because of a) substantial cross-loadings in the four-factor solution, and b) a factor with only two items in the six-factor solution.

Carrying on with the five-factor solution, items were removed iteratively based on cross-loadings (loadings higher than .32 on two or more items). First, item RH4 was removed because it loaded weakly on factors 3, 4, and 5 with -.31, .31, and .29, respectively. The analysis was rerun with the remaining 30 items. Item RM4 was removed next as it loaded weakly on factors 3, 4, and 5 with .38, -.34, and .37, respectively. After rerunning the model, item ES3 was removed as it loaded similarly on factors 2 and 3 with .36 and .41, respectively. Based on the factor solution with the remaining 28 items, item EP3 was removed as it had a loading of only -.30 on one factor (factor 3), and because it was not conceptually in line with that factor, which related to *differential value*. Finally, a satisfactory factor structure was obtained in which all remaining 27 items loaded significantly and higher than .32 on a main factor and had no cross-loadings higher than .32. It should be noted that some items loaded only weakly on their main factor (below .4). These items were retained as the questionnaire was still in its early development phase, but they were potential candidates for deletion in the confirmatory factor analysis.

The first factor comprised items relating to *unsophisticated task differentiation* as initially conceptualised and was thus labelled accordingly. The second factor comprised items relating to *performance over process recognition* as well as *social comparison*, which were initially thought to form two distinct dimensions relating to *evaluation*. Closely inspecting the content showed that most items related to evaluating students in front of others, which is why I labelled this factor *public evaluation*. Only the item EP5 ("My teacher gives us feedback about our progress, for example when we improve or get worse") did not fit well conceptually because it does not contain a public component. This item also had the lowest factor loading. Removing this item increased conceptual clarity and did not substantially affect the reliability of the scale (from $\alpha = .72$ to $\alpha = .71$). The third factor comprised items relating to *promotion of self-regulation* as initially conceived and was thus labelled accordingly. The fourth factor comprised items relating to *differential value* as originally conceptualised and was thus also named accordingly. Yet, item RM3 ("My teacher looks annoyed or disappointed when we make mistakes") also loaded on this factor. It could be argued that this item assesses differential value because students who experience teachers' annoyance and disappointment when making mistakes will likely feel less valued than students who make fewer mistakes. However, removing the item brought greater conceptual clarity and did not change the reliability of the scale (from $\alpha = .87$ to $\alpha = .87$). Lastly, items relating to *encouragement of help-seeking* and *handling of mistakes* formed one rather than two factors. Both aspects relate to *promotion of risk-taking*, which was hence used as the label. The ESEM model was rerun with the remaining 25 items. The fit indices for this model were good, with χ^2 (185) = 230.99, p = .01, CFI = 0.987, TLI = 0.979, RMSEA = 0.025 [0.012; 0.034], and SRMR = 0.027. The factor loading matrix is shown in Table 5.1.

Table 5.1

Results for the Exploratory Factor Analysis of Fixed-Ability Practices for Five Factors: Standardised Loadings and Factor Correlations

Factor	Item ID	1	2	3	4	5
1. Unsophisticated task differentiation	1. TD1	**.83**				
	2. TD2	**.72**				
	3. TD3	**.53**				
	4. TD4	**.76**				
2. Public evaluation	20. EP2		**.52**			
	22. EP4		**.63**			
	24. ES1		**.62**			
	25. ES2		**.44**			
	27. ES4		**.64**			
3. Promotion of self-regulation	34. ASR2			**.39**		
	35. ASR3			**.52**		
	36. ASR4			**.54**		
	37. ASR5			**.78**		
4. Differential value	38. DV1				**.65**	
	39. DV2				**.50**	
	40. DV3				**.67**	
	41. DV4				**.84**	
	42. DV5				**.69**	
5. Promotion of risk-taking	19. EP1					**.33**
	44. RH2					**.67**
	45. RH3					**.38**
	47. RM1					**.73**
	48. RM2					**.57**
	51. RM5r					**.36**
	52. RM6					**.56**
		1	2	3	4	5
1. Unsophisticated task differentiation		-				
2. Public evaluation		**.26**	-			
3. Promotion of self-regulation		.06	.13	-		
4. Differential value		**.36**	**.25**	**-.29**	-	
5. Promotion of risk-taking		-.09	.11	**.46**	**-.52**	-

Note. N = 411. Item RM5r was reverse scored. Factors 3 and 5 reflect adaptive practices. Loadings <.32 are suppressed. All main item loadings (in bold) were statistically significant at $p < .01$.

5.2 Results of confirmatory ESEM

5.2.1 First-order confirmatory ESEM

A first-order ESEM model using target rotation was estimated to test the student-level factor structure obtained in the exploratory analysis. The fit indices for this model were good, with χ^2 (185) = 226.19, p = .02, CFI = 0.990, TLI = 0.984, RMSEA = 0.023 [0.010; 0.033], and SRMR = 0.024, and the factor structure identified in the first split-half sample was replicated for the most part. However, items RH3 and RM5r[14] loaded only weakly on their respective main factors (.31 and .33, respectively), while item RH3 also loaded weakly on factor 3 (.29), and item RM5r also loaded weakly on factor 4 (-.32). Accordingly, these items were removed, and the analysis was rerun with the remaining 23 items. This constitutes the final student-level first-order model. Fit indices for this ESEM model are provided in Table 5.2 and final factor loadings can be found in Appendix G (see Table G.1).

Table 5.2

Final ESEM Model Fit Indices for Split-Half Sample 2

Model	χ^2	df	p	CFI	TLI	RMSEA	SRMR
ESEM	177.16	148	.05	0.993	0.988	0.022 [<0.001; 0.033]	0.021
H ESEM	181.14	153	.06	0.993	0.988	0.021 [<0.001; 0.032]	0.022
B ESEM	151.59	130	.09	0.995	0.990	0.020 [<0.001; 0.032]	0.019

Note. CFI = comparative fit index; TLI = Tucker–Lewis Index; RMSEA = root mean square error of approximation (with 90% CI); ESEM = exploratory structural equation modelling; H = hierarchical model; B = bifactor model.

5.2.2 Hierarchical ESEM

To test whether there was a higher-order FAPs factor, a hierarchical ESEM (H ESEM) model was estimated in which the final first-order ESEM model with 23 items was re-expressed using CFA, and the five first-order factors were specified to define the higher-order factor HF (for factor loadings, see Appendix G, Table G.2). Several issues arose when estimating the H ESEM model. First, there was a large standardised negative residual variance of -10.14 for factor 4 (*differential value*), which could indicate model misspecification. In addition, this factor had a correlation of 3.34 with the higher-order factor. Also problematic was that all *promotion of self-regulation* (Factor 3) items also loaded on the *promotion of risk-taking factor* (Factor 5). As the negative residual variance was very large, common practices such as fixing the residual variance to zero or forcing it to be larger than zero did not seem a viable option. Allowing residual variance between the highly correlated first-order factors 4 and 5 would have eliminated the negative residual variance, but this then led to a standardised residual correlation of -1.05 between these two first-order factors. All of this suggests that the H ESEM model did not fit the data well. The fit indices reported in Table 5.2 should be regarded with caution due to the negative residual variance.

[14] The “r” indicates that this item was reverse scored.

5.2.3 Bifactor ESEM

A bifactor ESEM (B ESEM) model was estimated to test whether there was a general FAPs factor in addition to the specific facets. Fit indices demonstrated an adequate fit (see Table 5.2). The factor loadings indicated that six items were not very well represented by the general factor, with loadings ranging between -.01 and .31 for items TD1, TD3, TD4, EP2, ES2, and ASR5 (see Appendix G, Table G.3). All other items had a loading above .32. In particular, items from the *differential value* and *promotion of risk-taking* facets had a high loading on the general factor. This is also reflected in the fact that the item loadings of these two facets were substantially reduced after accounting for the general factor. Such items should only be allowed to load on the general factor in future models (Morin et al., 2016). Overall, it cannot be said that a general factor underlies most items.

5.2.4 Model comparison

Based on model fit indices alone, it seems that all three student-level models fit the data similarly well because the differences in fit indices were all below the suggested cut-off values. However, the hierarchical factor model was not suitable because of the large negative residual variance. And even though the fit indices were excellent for the bifactor ESEM model, not all items were represented well by the general factor. Hence, a more parsimonious solution was preferred, and the first-order factor model was used for subsequent analyses.

5.3 Results of two-level ESEM

This section turns to the final two-level factor structure of student-perceived FAPs in the entire student sample. The student-level first-order structure was confirmed again using confirmatory ESEM with target rotation. The results showed good model fit, with χ^2 (148) = 195.99, p = .01, CFI = 0.993, TLI = 0.987, RMSEA = 0.02 [0.011; 0.027], and SRMR = 0.018. Standardised item loadings ranged from .38 to .91, and standardised factor correlations ranged from -.03 to -.63 (see Appendix G, Table G.4).

Two-level exploratory factor analyses with oblimin rotation were performed to assess the number of factors at the classroom level. Based on the results thus far, I estimated a series of models with five factors at the student level, and numbers of classroom-level factors from 1 to 5, as well as a model unrestricted at the classroom level. Fit indices for the six two-level factor models are presented in Table 5.3. Results showed that the model with only one classroom-level factor demonstrated a good fit, similar to the fit of the unrestricted model.

However, the SRMRb (SRMR for the between level) shows poor fit with SRMRb = 0.25 compared to SRMRb < .001 for the unrestricted model. While this fit index improved with a higher number of classroom-level factors, other fit indices did not show substantially better fit. The improved fit could be merely due to the increased number of parameters (Hooper et al., 2008). Moreover, models with more than one classroom-level factor were characterised by substantial cross-factor loadings and/or non-substantial loadings on postulated classroom-level factors. Thus, considering interpretability and parsimony, I decided in favour of the model with five student-level factors and one classroom-level factor.

Table 5.3

Results from the Two-Level Exploratory Factor Analyses of Fixed-Ability Practices Items: Model Fit of Models with Five Within-Level Factors and Varying Numbers of Between-Level Factors

Between-level factors	χ^2	df	p	CFI	TLI	RMSEA	SRMRw	SRMRb
1	470.10	378	<.001	0.980	0.973	0.017 [0.011; 0.022]	0.023	0.250
2	430.10	356	.004	0.984	0.977	0.016 [0.009; 0.021]	0.023	0.175
3	375.53	335	.063	0.991	0.986	0.012 [<0.001; 0.018]	0.023	0.082
4	354.48	315	.062	0.991	0.986	0.012 [<0.001; 0.019]	0.023	0.063
5	336.59	296	.052	0.991	0.985	0.013 [<0.001; 0.019]	0.023	0.049
Unrestricted	192.32	148	<.001	0.990	0.966	0.019 [0.010; 0.026]	0.023	<0.001

Note. N = 831. CFI = comparative fit index; TLI = Tucker–Lewis Index; RMSEA = root mean square error of approximation (with 90% CI); SRMRw = standardised root mean residual (within level); SRMRb = standardised root mean residual between (between level).

The factor loadings are shown in Table 5.4. Classroom-level factor loadings ranged from -.14 to 1.04. Some items, namely, TD1, TD4, EP2, EP4, and ES2, from the *unsophisticated task differentiation* and *public evaluation* dimensions loaded below .32 on the classroom-level factor and were thus not represented well. The three items with particularly low loadings, EP2, EP4 and ES2, were not included in the aggregated FAPs variable in later analyses. Accordingly, *public evaluation* constituted only a very small part of the uniform factor. One explanation might be that teachers who engaged in public evaluation did not necessarily also engage in some of the other FAPs. This is further addressed in the Discussion (see Section 11.1).

Table 5.4

Results from Two-Level Exploratory Factor Analysis of Fixed-Ability Practices Items: Within- and Between-Level Loadings for a Model with Five Student-Level Factors and One Classroom-Level Factor

Factor	Item ID	Within-level factor					Between-level factor
		1	2	3	4	5	1
1. Unsophisticated task differentiation	1. TD1	**.78**					.29
	2. TD2	**.65**					**.68**
	3. TD3	**.48**					**.31**
	4. TD4	**.72**					.29
2. Public evaluation	20. EP2		**.47**				.14
	22. EP4		**.52**				-.14
	24. ES1		**.73**				**.50**
	25. ES2		**.50**				.18
	27. ES4		**.64**				**.35**
3. Promotion of self-regulation	34. ASR2			**.32**			**-.64**
	35. ASR3			**.37**			**-.68**
	36. ASR4			**.50**			**-.55**
	37. ASR5			**.85**			**-.73**
4. Differential value	38. DV1				**.59**		**1.00**
	39. DV2				**.43**		**.92**
	40. DV3				**.78**		**.98**
	41. DV4				**.88**		**1.04**
	42. DV5				**.73**		**.97**
5. Promotion of risk-taking	19. EP1					**.36**	**-.80**
	44. RH2					**.67**	**-.89**
	47. RM1					**.82**	**-.74**
	48. RM2					**.48**	**-.87**
	52. RM6					**.50**	**-.66**
Standardised factor correlations							
1. Unsophisticated task differentiation		-					
2. Public evaluation		**.40**	-				
3. Promotion of self-regulation		**.10**	**.20**	-			
4. Differential value		**.39**	**.46**	**-.22**	-		
5. Promotion of risk-taking		-.06	-.08	**.48**	**-.54**	-	

Note. N = 831. Factors 3 and 5 reflect adaptive practices. Loadings <.32 are suppressed. All main item loadings (in bold) were statistically significant at $p < .01$.

5.4 Psychometric properties of student-perceived fixed-ability practices scales

This section addresses the psychometric properties of the scales assessing student-perceived FAPs. Table 5.5 presents the final scales resulting from factor analysis. The scales unsophisticated task differentiation, public evaluation, and differential value directly assess practices that were thought to be based on a fixed-ability notion. In contrast, the scales promotion of self-regulation and promotion of risk-taking assess practices that can be said to be based on the notion of a malleable ability. These scales were now reverse scored, so that all dimensions were coded in the same direction, with higher scores reflecting more frequent use of fixed-ability practices. Descriptive statistics are shown in Table 5.6. On average, students reported relatively little use of FAPs, in particular differential value. Public evaluation and little promotion of self-regulation were more common compared to the other practices. All scales had an adequate internal reliability, with $.70 < \alpha > .84$. In addition, the intraclass correlation coefficient (ICC) 2 indicated adequate reliability of the classroom mean ratings for all dimensions. Values above .70 are usually regarded as acceptable (Lüdtke et al., 2009). Lastly, the ICC 1 shows the proportion of between classroom variance on the total variance, which ranged between .13 and .25. This shows the extent to which ratings differ due to group membership, and results were similar to studies looking at comparable practices (e.g., Lüftenegger et al., 2017; Steuer et al., 2013).

Table 5.5

Scales and Items of the Final Fixed-Ability Practices Questionnaire for Students

Dimension	Item ID	Item
Unsophisticated task differentiation	1. TD1	My teacher gives easier tasks to pupils who don't do well in maths.
	2. TD2	My teacher gives less work to pupils who don't do well in maths.
	3. TD3	Pupils who do well in my maths class get harder tasks than others.
	4. TD4	Pupils who struggle in maths get fewer tasks to work on.
Public evaluation	20. EP2	My teacher points out pupils who work very hard in class
	22. EP4	My teacher points out when someone has improved in maths.
	24. ES1	My teacher points out pupils who get good maths grades, as an example to all of us.
	25. ES2	My teacher tells us which pupils get the highest scores on a maths test.
	27. ES4	My teacher talks about how well pupils do in maths in front of the whole class.
Little promotion of self-regulation	34. ASR2	My teacher asks us to think about the areas in maths that we need to improve. (reversed)
	35. ASR3	My teacher asks us to set our own learning goals. (reversed)
	36. ASR4	My teacher asks us to evaluate how well we are learning. (reversed)
	37. ASR5	My teacher asks us to evaluate our progress in maths. (reversed)
Differential value	38. DV1	My teacher seems to listen more carefully to pupils who do well in maths.
	39. DV2	My teacher asks pupils who do well in maths to give answers in front of the class more often than others.
	40. DV3	My teacher spends more time with pupils who are good at maths than with others.
	41. DV4	My teacher shows more interest in pupils who are good at maths than in others.
	42. DV5	My teacher gives less attention to pupils who don't do well in class.
Little promotion of risk-taking	19. EP1	My teacher praises us for trying hard in maths even if we don't do everything correctly. (reversed)
	44. RH2	My teacher answers questions carefully and thoroughly. (reversed)
	47. RM1	When someone makes a mistake, my teacher shows us how to learn from it. (reversed)
	48. RM2	When we make mistakes, my teacher gives us time to try again. (reversed)
	52. RM6	When we make a mistake, my teacher stops others from making fun of it. (reversed)

Table 5.6

Descriptive Statistics for Student-Reported Fixed-Ability Practices

Scale	N items	Range	*M*	*SD*	Skewness	α	ICC 1	ICC 2
Unsophisticated task differentiation	4	0-4	1.09	0.81	.62	.73	.25	.87
Public evaluation	5	0-4	1.54	0.84	.39	.70	.23	.85
Little promotion of self-regulation	4	0-4	1.94	0.89	.25	.72	.18	.80
Differential value	5	0-4	1.05	0.90	.96	.84	.18	.81
Little promotion of risk-taking	5	0-4	1.12	0.79	.83	.78	.13	.75

Note. N = 831. ICC = Intraclass correlation coefficient (1 = proportion of between classroom variance on total variance; 2 =reliability of the classroom mean ratings).
All skewness statistics are significant at .01.

5.5 Teacher-reported fixed-ability practices

As both students and teachers rated teachers' FAPs, I also sought to establish whether the teacher scales had adequate psychometric properties. Due to the small teacher sample size (only 26 teachers completed the FAPs questionnaire), it was not possible to perform factor analysis. Hence, the factor structure obtained in the student sample was used to form the teacher scales. Descriptive statistics are provided in Table 5.7. Internal reliability of the five scales was assessed using Cronbach's alpha. Notably, the sample size was also quite small for alpha, which is why bootstrapping was used to estimate 95% confidence intervals. Analyses were performed in R with the MBESS package (Kelley, 2017) using the bias-corrected and accelerated bootstrap with 1000 samples. Results clearly showed that adequate reliability for the teacher scales was not reached as the point estimates were below the common cutoff of .70, and the widths of the 95% confidence intervals indicated high uncertainty of the estimates. I subsequently examined whether adequate reliability could be reached when removing individual items from the scales using the *alpha if item deleted* statistic in SPSS. None of the scales reached and alpha of .70 when deleting any of the items. As a consequence, teacher-reported practices were not considered in the examination of relationships between teachers' motivational beliefs and FAPs or between FAPs and students' motivational beliefs and achievement.

Table 5.7

Descriptive Statistics for Teacher-Reported Fixed-Ability Practices

Scale	Range	*M*	*SD*	α [95% CI]
Unsophisticated task differentiation	1.00 – 3.00	1.87	0.56	.62 [.38; .78]
Public evaluation	1.20 – 3.00	2.08	0.48	.40 [.00; .64]
Little promotion of self-regulation	1.00 – 2.75	1.68	0.48	.56 [.27; .82]
Differential value	0.00 – 1.60	0.97	0.38	.59 [.33; .81]
Little promotion of risk-taking	0.20 – 1.20	0.70	0.30	.54 [.30; .76]

Note. N = 26 for all variables except self-regulation where N = 25 because of a missing value.

5.6 Brief summary on RQ 1

This chapter aimed to answer the question "What are fixed-ability practices?" Analyses have shown that student-perceived FAPs can be characterised by five distinct dimensions at the student level and one dimension at the aggregated classroom level. The individual dimensions are *unsophisticated task differentiation, public evaluation, little promotion of self-regulation, differential value,* and *little promotion of risk-taking*. Although a bifactor model with a general FAPs factor fit as well as a model without such a factor, the general factor did not represent all items well and was thus not used in further analyses. The student FAPs scales showed adequate psychometric properties, whereas the teacher scales were unreliable. Hence, only the student scales were used in further analyses.

Chapter 6
Analysis Approach for Quantitative Relationships (RQs 2 and 3)

After having established the dimensionality of fixed-ability practices (FAPs), at least as perceived by students, this chapter presents the analyses for the relationships between teachers' motivational beliefs and their engagement in student-perceived FAPs (research question (RQ) 2) as well as the longitudinal relationships between student-perceived FAPs and students' motivational beliefs and achievement (RQ 3). This chapter begins by outlining the analytical approach in more detail (Section 6.1). Then, the data preparation and preliminary analyses are described (Section 6.2). This is followed by a description of the main analyses (Section 6.3).

6.1 Analyses overview

To address RQ 2, that is, to examine whether teachers' motivational beliefs predict their engagement in FAPs, I initially planned to conduct regression analyses with teachers' beliefs as the predictors and teacher-reported FAPs as the outcome, as well as regressions with teachers' beliefs as the predictors and students' shared perceptions of FAPs (i.e., the uniform FAPs dimension) as the outcome. Teacher-reported FAPs could not be used in these analyses because the scales were not reliable as described in the previous chapter. Correlation analyses later showed that none of the relationships between teachers' motivational beliefs and students' shared perceptions of FAPs was significant (see Section 7.1.1). Thus, conducting regression analyses was not justified, and there was no need to present these planned analyses here.

To answer RQ 3, I examined whether student-perceived FAPs (but not teacher-reported FAPs as initially planned) predict students' motivational beliefs and achievement. I controlled for initial beliefs and achievement because this provides stronger evidence for a causal relationship. Accordingly, I first established longitudinal measurement invariance for the motivational beliefs variables at Time 1 and Time 3 to ascertain that the meaning of the constructs did not change. This was done in Mplus by estimating factor models with increasingly restrictive equality constraints across the time points and comparing fit indices for these models.

Subsequently, I ran a series of multivariate two-level regression models to predict beliefs and achievement with the five FAPs dimensions at the student level and the aggregated uniform FAPs dimension at the classroom and teacher level. Due to the complex data structure described earlier (see Section 4.1.2), a cross-classified multiple membership (CCMM) model was used. This model takes the proportion of time that students were exposed to different teaching practices into account by specifying that students belonged to multiple teachers and classrooms. The models are explained in more detail below. The analyses were performed using Bayesian statistics and were conducted with the brms package in R (Bürkner, 2017, 2018). Using Bayesian rather than frequentist statistics had pragmatic reasons as, to the best of my knowledge, no software yet existed which could estimate such complex models taking multiple membership into account. While there are some philosophical debates about the superiority of the two approaches, from a methodological perspective, there seems to be an understanding that both are valid and can make important contributions (Berger & Bayarri, 2004). Accordingly, Berger and Bayarri (2004) argued that both should be used, and I concentrated on the pragmatic benefits that the flexibility of Bayesian statistics offers for estimating complex models (Gelman et al., 2013).

Notably, an important difference between frequentist and Bayesian statistics is that the former is mostly interested in point estimates, whereas Bayesian statistics is interested in probability distributions of parameters, called the posterior probability distribution (Van de Schoot et al., 2014). This distribution quantifies how probable it is for a given parameter to lie in a certain region given prior knowledge about the parameter and the likelihood (i.e., the data). The parameter is then usually described by the mean of that posterior distribution accompanied by the 95% credible interval, a density interval. In contrast to frequentists' confidence intervals, credible intervals directly quantify the probability that a parameter lies within the limits of the interval (Van de Schoot et al., 2014).

6.2 Data preparation and preliminary analyses

Several steps were necessary to prepare the data before conducting the main analyses. This included the formation of composite scores for variables measured with multiple items, standardisation of achievement scores, and the assignment of weights for the CCMM models. I also sought to establish longitudinal measurement invariance for the motivational beliefs variables.

6.2.1 Composite scores

It would have been ideal to use latent variable modelling rather than treating the constructs of interest as manifest variables because the former controls for measurement error. However, this was not possible because of the relatively small sample size at the classroom and teacher level, which would not give a lot of power to an analysis that includes all indicator items. Moreover, I could not find any software (e.g., Mplus, MLWin) or R package that could yet handle CCMM models using latent variables. Accordingly, composite scores were created for each variable by averaging across the items constituting a scale. For the aggregated, uniform FAPs dimension, three items (EP2, EP4, and ES2) from the public evaluation scale were excluded from the composite score because they had very low factor loadings as described in the previous chapter. Thus, 20 variables were used to form the overall FAPs dimension.

6.2.2 Standardised achievement scores

While the Key Stage 2 assessment at the end of Year 6 is a statewide test so that the maths performance of all students could easily be compared, students did not take a standardised test at the end of Year 7. Accordingly, tests differed between schools, and performance across schools could not be compared. Within schools, most but not all classes took identical tests. Test performance for both the Key Stage 2 and end of Year 7 tests was standardised across all students taking the same end of year test, so that the reference group against which students' performance was compared remained the same.

6.2.3 Weights for group membership

Multiple membership models allow that weights are assigned to different group memberships for each individual to reflect the proportion of impact that a given group is thought to have on an individual. There are several approaches to assigning weights (see Fielding & Goldstein, 2006). One approach is to assign equal weights to every teacher and classroom that a student belonged to irrespective of the length of membership. Another approach is to assign weights based on the proportion of time a student belonged to a given group, which implies that the groups a student belonged to for longer have a greater influence. A third option is to weight more recent clusters more heavily than clusters a student belonged to earlier in the year. There are of course many other factors that could determine the strength of influence a teacher and class could have. In this study, one could argue that earlier teachers have a greater influence because students just transitioned to a new school so that earlier teachers could have set the new tone. The quality of the relationships that students have with their teachers and peers might also determine the strength of influence. Given the absence of such information, it seemed most reasonable to assign weights based on the proportion of time a student was taught by a given teacher and belonged to a given class. This is also the recommended approach by Goldstein et al. (2007).

It should be noted that a simulation study showed that the weighting scheme appears to have relatively little influence on results as long as the multiple membership structure is not completely being ignored (Wolff Smith & Beretvas, 2014). There were some irregularities in the applied weighting scheme. For instance, some classes were taught by some teachers for only a few weeks and no data about their teaching practices was collected. More details on these irregularities can be found in Appendix H (H.1).

6.2.4 Longitudinal invariance testing

Establishing longitudinal measurement invariance requires the estimation of a series of factor models with increasingly restrictive equality constraints across the time points and comparisons of fit indices for these models. Three models were estimated: a) a configural model to test whether the same general pattern of factor loadings held across time, b) a scalar model to test whether the factor loadings and thresholds held across time, and c) a residual invariance model to test whether residual variances held across time. Note that I went straight from the configural to the scalar model rather than first fitting one model with constrained loadings but freed thresholds. This is the recommended approach (Muthén & Muthén, 1998–2017) because loadings and thresholds jointly define item functioning in ordinal models (Sass, 2011). When residual invariance holds, all differences in the items between time points can be attributed to time differences in the common factors alone; hence, comparisons of mean differences are valid (Chen, 2007; Liu et al., 2017).

ESEM models were estimated in Mplus with the imputed data set including all students who completed both Time 1 and Time 3 questionnaires. The complex option was used to account for grouping in classes. To the best of my knowledge, there was not yet a software in which measurement invariance could be examined while taking the multiple membership structure into account. Hence, I used class membership at Time 3 as the cluster variable because this was usually the classroom that students belonged to for the majority of the time. Moreover, Time 1 motivational beliefs were assessed at the beginning of the school year, so that group membership was not thought to have notably influenced students' beliefs at that time. Models were estimated using the WLSMV estimator, Theta parameterisation, and oblique target rotation. Note that the attribution item was not included as it was a single item, which cannot form a factor. Corresponding item residuals were correlated across time points, which is a distinctive feature of longitudinal measurement invariance testing (Liu et al., 2017).

First, I assessed the model fit of the configural model. Factor loadings and thresholds were allowed to differ across time, residual variances were fixed at one, and factor means were fixed at zero for both time points. The metric of the factor was set by fixing the factor variance to one at both time points. Second, I fitted the scalar model in which the factor loadings and thresholds were constrained across time. Residual variances were fixed at Time 1 and freed at Time 3, and factor means were fixed at zero at Time 1 and freed at Time 3. The metric of the factors was set by fixing the factor variance at one for Time 1 and freeing it for Time 3. Lastly, I fitted the residual variance model in which residual variances were fixed at one at both time points in addition to the constraints of the scalar model.

If the more restrictive model fits worse than the previous less restrictive model, the assumption of measurement invariance should be rejected. It is advised to use multiple indicators to evaluate whether the decrement in fit is within acceptable ranges (Chen, 2007; Cheung & Rensvold, 2002; Sass et al., 2014). Hence, I examined changes in RMSEA (ΔRMSEA), CFI (ΔCFI), TLI (ΔTLI), and SRMR (ΔSRMR). According to the cutoff criteria proposed by Chen (2007), scalar and residual invariance are supported if $\Delta\text{RMSEA} < 0.015$, $\Delta\text{CFI} < 0.01$, and $\Delta\text{SRMR} < 0.01$. Sass et al. (2014) recommend that a $\Delta\text{TLI} \leq -0.01$ can be taken as evidence for measurement invariance, although Marsh et al. (2010) argued that a ΔTLI close to zero might be more appropriate.

6.3 Examining the effect of student-perceived fixed-ability practices on students’ motivational beliefs and achievement

6.3.1 Overview of the cross-classified multiple membership models

To estimate the effects of student-perceived FAPs on students’ motivational beliefs and achievement, I estimated two-level cross-classified multivariate regression models in which students (Level 1) were nested in classrooms (Level 2) and teachers (Level 2). Every outcome variable was predicted by its own set of predictors, whereby the group-level effects were modelled as correlated across regression models. At the student level, every outcome was predicted by sex (girl), the lagged outcome variable[15], as well as the five FAPs dimensions. At the classroom/teacher level, the outcomes were predicted by the uniform FAPs dimension,

[15] Lagged outcome variables are the same variable as the outcome but measured at Time 1 rather than Time 3.

which was the aggregated score. Note that because students reported practices for only one of their teachers, the effects of the five student-level FAPs dimensions could only be examined using a sample with students who were taught by a single teacher throughout the year. For this sample, henceforth called the "single-teacher sample," I estimated a cross-classified (CC) model rather than a CCMM model. The effects of the uniform FAPs dimension at the second level were examined for the entire sample, including students who belonged to multiple teachers and/or classes, using a CCMM model. This model did not include the individual FAPs dimensions at the student level. For the single-teacher sample as well as the entire sample, I first estimated the simplest possible model, the unconditional "null" model or random intercept model, which did not include any predictors. This was followed by the full model including all first-level predictors as random effects, and the uniform FAPs dimension at the second level as a fixed effect.

Accordingly, four multivariate models (CC_{Null}, $CCMM_{Null}$, CC, CCMM) were estimated, each with the outcome variables fixed mindset, mastery goal orientation, performance-approach goal orientation, performance-avoidance goal orientation, failure-to-ability attribution, self-concept, and achievement. First, I estimated a two-level cross-classified random intercept model for the single-teacher sample (CC_{Null}). Second, I estimated a two-level cross-classified multiple membership random intercept model for the entire sample ($CCMM_{Null}$). Third, I estimated a two-level cross-classified random effects model for the single-teacher sample, including sex, the lagged outcome, and the five FAPs dimensions at the student level and the uniform FAPs dimension at the teacher/classroom level (CC). Fourth, I estimated a two-level cross-classified multiple membership random effects model with the entire sample, including sex (girl) and the lagged outcome at the student level and the uniform FAPs dimension at the classroom/teacher level (CCMM). All model formulae are presented in Appendix H (H.2). I will now describe the model fitting choices including the selection of the distributions that were used to model variables, the inclusion of random effects, the centring approach for predictors, the selection of priors, and model estimation.

6.3.2 Model fitting choices

6.3.2.1 Distribution family

One needs to decide how to model the conditional distribution of the outcome variables, called the distribution family (Bürkner & Vuorre, 2019). The distribution family represents how the outcome values are distributed for specific values of the predictor variables. I used the Skew_Normal family in brms for all composite variables because almost all student variables were significantly skewed. This distribution can model both right- and left-skewed data and is described by three parameters: one for location (mu), scale (sigma), and skewness (alpha). For predicting failure-to-ability attribution, which was a single item, I used the cumulative family with a probit link[16]. The cumulative family is appropriate for ordinal variables that are assumed to have a continuous underlying latent variable. The probit link uses the inverse Gaussian transformation, which leads to ordered probit regression (Bürkner & Vuorre, 2019). I modelled the ordinal predictor variable failure-to-ability attribution as a monotonic effect, which assumes that the outcome has a monotonically increasing or decreasing relationship with the predictor (see Bürkner & Charpentier, 2020). That way, adjacent predictor categories are not assumed to have the same distance, but distances are directly estimated from the data.

6.3.2.2 Inclusion of random effects

The effect of a predictor on the outcome can be modelled as fixed or as random. Both approaches aim to control for clustering and the differing relationships between predictor and outcome variables within and between clusters (Bell et al., 2019). When one uses fixed effects, the effects of a predictor on the outcome are assumed to be equal across all clusters, whereas modelling predictors as random allows that their effects vary across clusters. Importantly, fixed effects models assume that the relationship between predictor and outcome in a given cluster are completely independent of each other. If the cluster unit is *classroom*, it would be assumed that the relationship between predictor and outcome in one classroom cannot provide any information about the same relationship in a different classroom (Bell et al., 2019). The fixed effects coefficients then simply represent the average effect across all classrooms. In contrast, random effects are drawn from a common, normal distribution (Bell et al., 2019), and information about the effects in one cluster are used when estimating the effects in other clusters, which is referred to as partial pooling (McElreath, 2016). More specifically, the random effects estimates are shrunk towards the mean, whereby estimates that are less reliable (e.g., estimates obtained from classrooms with only a few students) or very extreme are shrunk the most. This leads to more accurate estimates of cluster effect means, and the fixed effects in such models then present population regression coefficients, which is what one is ultimately interested in (McElreath, 2016). The random effect estimates show the differences in predictor-outcome relationships between clusters. Accordingly, random effects are expressed in variance around the average regression coefficient (Snijders,

[16] A link function relates the expected value of the outcome variable to the linear predictors in the model.

2005). Bell et al. (2019) argue that random effects models provide all the information that fixed effects models provide but even more, which is why they regard them as the superior approach. Hence, I estimated random effects models.

In addition to the five student-level FAPs dimensions, the effects of the student-level control variables sex and lagged outcomes were also modelled as random. Failing to account for existing varying effects of such variables can bias the within-cluster effects estimates, which can also reduce the precision of the estimated context effects (Heisig et al., 2017). Moreover, Bell et al. (2019) found in a simulation study that even if the normality assumption for random effects is violated, models have only modest bias, while not including random effects leads to anti-conservative standard errors.

6.3.2.3 Centring of predictors

An important decision in multilevel modelling regards how to centre the predictor variables. Centring is paramount because it helps to distinguish group-level effects from individual-level effects. There are generally two approaches: centring within clusters (CWC) using cluster means and centring within the whole sample using the grand-mean (CGM). The centring option changes the interpretation of the regression coefficients and can even affect the significance level of effects (Lüdtke et al., 2009). The main difference between the approaches lies in the interpretation of the regression coefficients at the group level. When CGM is used for student ratings of a teaching practice, individual differences in students' perceptions of this practice are controlled for when assessing the effect of the aggregated, shared perceptions (Lüdtke et al., 2009). CGM thus results in a partial regression coefficient at the group level. When CWC is used, the between-class effect of the aggregated perceptions is not controlled for interindividual differences in student perceptions. Lüdtke et al. (2009) argue that CGM might be inappropriate when students rate a class-level construct, here FAPs, that is assumed to cause students' perceptions because differences in individual student perceptions can already be an effect of the construct that is being assessed by them. Thus, controlling for individual differences can be argued to eliminate a crucial part of the aggregated perceptions. Hence, CWC was used for the five student-level FAPs dimensions, which allowed me to examine the effect of the aggregated and individual perceptions on student outcomes, whereby individual differences were not controlled for when examining the aggregated effects. I used CGM for the student-level control variables sex and lagged outcomes to control for these variables when estimating the effects of the aggregated, uniform FAPs dimension. CWC would defy the purpose of control variables since the aggregated

FAPs dimension would be evaluated independently of them (Enders & Tofighi, 2007; Lüdtke et al., 2009).

The level-two predictor can only be centred using CGM or not be centred at all because all students in the same class with the same teacher have the same value for the uniform FAPs variable. CGM was preferred to ease the interpretation of intercepts. With CGM, the intercept is interpreted as the expected outcome for a student in classroom c and teacher t who is at the overall mean of the predictor variable, that is, the mean across all students in the sample. With CWC, the intercept is interpreted as the expected outcome for a student in classroom c and teacher t whose predictor values are equal to the mean across students in classroom c and teacher t.

6.3.2.4 Selection of priors

One vital step in Bayesian analysis is setting the priors for all estimated parameters. Priors are probability distributions that show what values a given parameter can take and how much uncertainty there is (Depaoli & Van de Schoot, 2017). Priors and likelihoods (the data) are combined in the analysis to form the posterior distribution - the estimates and credible intervals of the parameters. Generally, one distinguishes between non-informative, weakly-informative, and informative priors. How informative a given prior is always depends on the metric and scale of the parameter. The sample size is also important as priors have less influence in larger samples because the likelihood will overpower the prior. Non-informative priors represent a complete lack of any prior information about the parameter and are typically flat (i.e., they assign approximately equal probability to all values). Another example is a normal distribution with a mean of zero and a very large standard deviation, such as 1000. While these priors are common in practice, many researchers warn against using uninformative priors because they can pull the posterior distribution towards very extreme, very unlikely, or even impossible values (Gelman et al., 2017; Lambert et al., 2005). Weakly-informative priors introduce some prior knowledge, which usually helps to stabilise the estimation and to prevent inappropriate inference because they cover the range of possible values but do not provide enough information to pull the posterior away from the likelihood (Gelman et al., 2008). Lastly, informative priors provide strict information about the possible parameter values because they assign little uncertainty to a given estimate.

According to Gelman et al. (2017), choosing the prior should be guided by the question of whether the prior distribution can generate the data one expects to see. This means that the majority of the prior distributions' density should contain the range of parameter values that are reasonable for a given model. Similarly, Schad et al. (2019) recommend setting priors that contain just enough information to exclude implausible values while still allowing extreme values. This principle guided me in the prior selection. All priors used in the models can be found in Appendix H (H.3). For most parameters, I decided to use the brms default priors. These are very weakly informative but provide just enough information to regularise and stabilise the estimation, while having minimal influence on the estimation (see https://rdrr.io/cran/brms/man/set_prior.html). I used these for the standard deviations of group-level (i.e., random) effects and the effects specific to the skew_normal family, namely alpha (the skew of outcome variables) and sigma (the standard deviation of the outcome variables).

For fixed effect parameters (regression coefficients and intercepts), the brms default priors are improper flat priors over the reals, which means that all values have an equal probability. For this prior class, I decided to choose custom priors to exclude parameter values that are extremely unlikely, if not to say impossible, while still allowing extreme values. For regression coefficients for the unstandardised, continuous outcomes, I set a normal prior with a mean of zero and standard deviation of two (N(0,2)). For their intercepts, I set a prior of N(2,2). For standardised achievement, I set a prior of N(0,1) for fixed effects and intercepts, which still allows for extreme values. For the attribution variable, which was modelled as a monotonic effect, I also set a prior of N(0,1) for fixed effects, and used the brms default prior for the threshold. Here, the regression coefficient is the average change in the outcome when moving from one response category to the next. Before running analyses with these custom priors for fixed effects, I ran all analyses using the brms default priors. This is the recommended procedure by Gelman and Hill, who regard non-informative priors as provisional priors that should be used as a starting point but not for the final results (2007).

6.3.2.5 Model estimation

The brms package (via STAN) uses the Hamiltonian Monte Carlo sampler algorithm (MCMC) to estimate the posterior distribution. To estimate the models, I used four MCMC chains with 2000 iterations each. Of these, 1000 iterations per chain were warmups (i.e., initial iterations that are discarded per chain). Thus, the total post-warmup sample was 4000. These are the default settings of brms. Using three to four chains is the convention, which is usually enough to ensure appropriate sampling (McElreath, 2016). Regarding iterations, an effective sample size[17] of 200 is usually sufficient to get good estimates in most regression applications (McElreath, 2016). The initial starting value for the chains was set to zero because this helped to start the analyses.

6.3.3 Model checking

6.3.3.1 Model convergence

Before looking at the estimation outcomes, some essential checks need to be performed to ascertain that the MCMC chains have converged to the posterior distribution. First, one needs to visually inspect the trace plots (Gabry et al., 2019), which show the values a parameter takes over the runtime of the chains. The plot should show a rapid up-and-down variation with no continuing trends so that the plot at one point looks very similar to that at a different point. Next, there should be no divergent numerical trajectories (i.e., the leap-frog integrator diverges from the energy-conserving trajectory) because these call any estimates of the MCMC chain into question (Gabry et al., 2019). Divergent trajectories are automatically flagged by brms. Another statistic that assesses convergence is $\widehat{R}$, which indicates whether there is a significant difference between the between-chain variance and within-chain variance. Values very close to 1 indicate no problem, whereas values of 1.1 and above indicate problems with model convergence (Gelman & Rubin, 1992).

Moreover, the autocorrelation of samples returned by the MCMC algorithm should be examined. The autocorrelation between an iteration and the iteration k-steps before it in the same chain should decrease with increasing k. High degrees of autocorrelation can indicate problems in the estimation process (Depaoli & Van de Schoot, 2017; Gelman et al., 2013). Related to the correlation of iterations or samples is the measure of effective sample size. The higher the autocorrelation between samples, the lower the effective sample size compared to the actual one (i.e., the number of iterations after warm-up) because samples that are highly correlated provide less information than independent samples. If there is no autocorrelation, the effective sample size equals the actual sample size. The number of effective samples can be higher than the actual sample size using the NUTs algorithm in Stan for parameters which have a close to Gaussian posterior distribution and are little dependent on other parameters. If the effective sample size is less than 10% of the actual sample size, the estimations are generally considered to be problematic (see https://mc-stan.org/bayesplot/reference/MCMC-diagnostics.html). I checked these diagnostics using the R package ShinyStan, which

[17] This is the number of effectively independent draws from the posterior distribution that the Markov chain is equivalent to.

provides visual and numerical convergence diagnostics for MCMC simulations (Stan Development Team, 2018b).

6.3.3.2 Model fit: Posterior predictive checks

One can assess whether the model fits the data well if the model can be used to predict new values that resemble the observed data (Gabry et al., 2019). This can be checked by drawing new data sets from the posterior predictive distribution. For the continuous variables, that is all except failure-to-ability attribution, I compared Kernel density estimates of the observed dataset y with the density estimates for 100 simulated datasets y_{rep} as well as boxplots for ten replicated samples. Posterior predictive checks for failure-to-ability attribution were done using histograms. These checks were performed with the bayesplot package in R (Gabry & Mahr, 2020).

6.3.4 Interpretation of regression results

I based the interpretation of the regression results upon the mean of the parameter estimate and the 95% credible interval (CI) for that estimate. An effect was considered supported if the posterior distribution for the regression coefficient, represented by the 95% CI, did not include zero because the posterior probability of the coefficient having a non-zero value then is > .95. CIs that include zero do not provide enough support for a definitive judgement about an effect, but trends may be discerned. Reporting the posterior distribution was preferred over hypothesis testing using the Bayes factor as the latter approach is heavily debated amongst statisticians, mainly because its results are strongly influenced by the priors and because it is argued that the posterior distribution provides more information (for a discussion, see Tendeiro & Kiers, 2019). When looking at effects that were estimated in both the CC and CCMM models, I assumed that the CCMM model results are more reliable given the larger sample size and, thus, less error and greater representativeness.

Chapter 7
Results for Quantitative Relationships (RQs 2 and 3)

This chapter presents the results for the analyses described in the previous chapter. Thus, it addresses the following research questions (RQs):

2. Do teachers' motivational beliefs predict their engagement in student-perceived and/or teacher-reported fixed-ability practices? If so, which beliefs predict what practices?

3. Do student-perceived and/or teacher-reported fixed-ability practices predict students' motivational beliefs and achievement? If so, which practices predict what beliefs and which practices predict achievement?

As described in previous chapters, only student-perceived but not teacher-reported FAPs were used in analyses. First, preliminary results are presented including descriptive statistics and correlations for all variables, results of the longitudinal measurement invariance assessment for students' motivational beliefs, and model convergence and fit for the Bayesian cross-classified (CC) and cross-classified multiple membership (CCMM) models (Section 7.1). I will describe correlations between teachers' motivational beliefs and their engagement in student-perceived fixed-ability practices (FAPs) to address RQ 2. Subsequently, the relationships between student-perceived FAPs and students' motivational beliefs and achievement (RQ 3) are presented (Section 7.2). I will then briefly summarise results on RQs 2 and 3 (Section 7.3).

7.1 Preliminary results

7.1.1 Descriptive statistics and correlations

Descriptive statistics for all variables, including students' motivational beliefs at Time 1 and Time 3, teachers' motivational beliefs at Time 1, and student-perceived FAPs at Time 2, can be found in Table 7.1. Maths achievement is not presented as it is a standardised score. The reliability of each scale was examined using the internal consistency coefficient Cronbach's α. Reliability was considered acceptable if $\alpha > .70$ (Field, 2009). This was the case for all scales except for the performance-avoidance goal orientation scale for teachers, which had a reliability of only $\alpha = .36$. Since this scale had only three items, it was not possible to remove any items to increase reliability. Accordingly, any correlation with that variable should not be considered. Please note that the intraclass correlation coefficients (ICCs) for the FAPs scales have been discussed earlier (see Section 5.4).

Correlations between all student variables and teacher variables are shown in Table 7.2 and Table 7.3, respectively. Relevant for RQ 3 are relationships between FAPs and students' motivational beliefs. Some motivational beliefs variables at Time 1 correlated with student-perceived FAPs reported at Time 2, which indicates that students' initial beliefs may influence how they perceived teaching practices. This shows the importance of controlling for initial beliefs in the regression models. Importantly, compared to Time 1 beliefs, there were more significant and overall stronger correlations between student-perceived FAPs and motivational beliefs at Time 3. Please note that correlations do not present one-to-one relationships because students reported practices for only one teacher even if they were taught by multiple teachers, which presumably diluted correlations. The correlations were in the expected directions. The FAPs dimensions correlated negatively but not always significantly with achievement at both time points except for little promotion of self-regulation. Unexpectedly, little promotion of self-regulation was positively associated with achievement, which is further explored in the main analyses.

Table 7.4 presents the correlations between teachers' motivational beliefs and student-perceived FAPs at both the student and teacher level. Relevant to RQ 2 are the correlations between teachers' motivational beliefs and the aggregated, uniform FAPs dimension at the teacher level. These relationships must be examined at the teacher level as these are teacher-level constructs. None of the correlations between teachers' motivational beliefs and the uniform FAPs dimension was significant. Hence, correlations did not support that teachers' motivational beliefs predicted the teaching practices they engaged in according to students. The absence of significant correlations showed that it was not necessary to follow this up with a regression analysis. Moreover, the teacher sample was very small for conducting a multiple regression analysis, so that detecting any significant effects was highly unlikely. For instance, Hair et al. (2014, p. 170) argued that only a simple regression with a single independent variable can be conducted if the sample is smaller than 30, which is the case here.

Table 7.1

Scale Descriptive Statistics

Scale (*N* items)	Range		*M*		*SD*		Skewness		α		ICC 1	ICC 2
Student motivational beliefs	T1	T3	T1	T3	T1	T3	T1	T3	T1	T3		
Fixed mindset (4)	0.00 – 4.00	0.00 – 4.00	1.21	1.17	0.85	0.81	0.62**	0.69**	.73	.74	-	-
Mastery goal (5)	1.00 – 4.00	0.80 – 4.00	3.30	3.23	0.59	0.60	-0.93**	-1.02**	.76	.78	-	-
Performance-approach goal (5)	0.00 – 4.00	0.00 – 4.00	1.59	1.46	0.91	0.94	0.38**	0.48**	.84	.87	-	-
Performance-avoidance goal (4)	0.00 – 4.00	0.00 – 4.00	1.84	1.69	0.95	0.94	0.07	0.24**	.73	.76	-	-
Failure-to-ability attribution (1)	0.00 – 4.00	0.00 – 4.00	1.34	1.50	1.15	1.14	0.48**	0.32**	-	-	-	-
Self-concept (5)	0.00 – 4.00	0.00 – 4.00	2.82	2.70	0.74	0.76	-0.62**	-0.50**	.86	.88	-	-
Teacher motivational beliefs	T1		T1		T1		T1		T1			
Fixed mindset (4)	0.00 – 1.75		0.75		0.49		0.20		.76		-	-
Mastery goal (3)	2.67 – 4.00		3.52		0.46		-0.40		.81		-	-
Performance-approach goal (4)	0.50 – 3.75		1.77		0.77		0.63		.84		-	-
Performance-avoidance goal (3)	2.00 – 4.00		3.10		0.52		0.05		.36		-	-
Failure-to-ability attribution (1)	0.00 – 2.00		2.30		0.88		0.08		-		-	-
Perceived ability/expectation (2)	1.57 – 2.94		2.46		0.28		-1.11**		-		-	-
Student-perceived FAPs	T2		T2		T2		T2		T2		T2	T2
Unsophisticated task differentiation (4)	0.00 – 4.00		1.09		0.81		0.67**		.73		.25	.87
Public evaluation (5)	0.00 – 4.00		1.54		0.84		0.39**		.70		.23	.85
Little promotion of self-regulation (4)	0.00 – 4.00		1.94		0.89		0.25**		.72		.18	.80
Differential value (5)	0.00 – 4.00		1.05		0.90		0.96**		.84		.18	.81
Little promotion of risk-taking (5)	0.00 – 4.00		1.12		0.79		0.83**		.78		.13	.75

Note. Overall $N_{Student}$ = 927 (N_{T1} = 873; N_{T2} = 831, N_{T3} = 807), Overall $N_{Teacher}$ = 30. T = Time; ICC = Intraclass correlation coefficient (1 = proportion of between classroom variance on total variance; 2 =reliability of the classroom mean ratings). All scales range from 0 to 4.
*p<.05, **p<.01

Table 7.2

Pearson's Correlations Amongst Students' Motivational Beliefs and Student-Perceived Fixed-Ability Practices

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Beliefs T1																			
1. Fixed	-																		
2. Mastery	**-.30**	-																	
3. Approach	**.20**	**.16**	-																
4. Avoidance	**.23**	**.14**	**.67**	-															
5. Attribution	**.18**	*-.08*	-.02	.03	-														
6. Self-con.	**-.27**	**.34**	.18	.04	**-.35**	-													
Beliefs T3																			
7. Fixed	**.46**	**-.23**	**.12**	**.15**	**.24**	**-.24**	-												
8. Mastery	**-.25**	**.45**	.04	-.03	**-.11**	**.24**	**-.32**	-											
9. Approach	**.12**	**.12**	**.46**	**.37**	-.01	*.10*	**.17**	**.14**	-										
10. Avoidance	**.15**	*.09*	**.39**	**.45**	*.09*	.05	**.29**	.06	**.70**	-									
11. Attribution	**.17**	-.07	-.02	.03	**.32**	**-.35**	**.38**	**-.11**	.07	**.20**	-								
12. Self-con.	**-.19**	**.27**	**.18**	.06	**-.27**	**.63**	**-.39**	**.36**	**.19**	.02	**-.44**	-							
FAPs																			
13. Task	**.11**	-.03	.05	.05	*.08*	-.02	*.09*	-.04	**.12**	*.09*	**.15**	-.01	-						
14. Evaluation	**.15**	.03	**.10**	**.11**	*.09*	-.01	.05	-.04	**.14**	**.11**	.02	-.001	**.25**	-					
15. Self-reg.	.02	**-.15**	-.02	-.03	.02	.01	**.10**	**-.17**	.04	-.05	.02	-.07	*-.07*	-.07	-				
16. Value	**.17**	**-.16**	*.09*	**.14**	**.12**	-.06	**.19**	**-.19**	**.19**	**.17**	**.17**	**-.15**	**.37**	**.39**	**.30**	-			
17. Risk	**.11**	**-.20**	.01	.05	.05	*-.08*	**.18**	**-.28**	**.12**	**.12**	**.11**	**-.17**	*.09*	.06	**.53**	**.54**	-		
Achievement																			
18. Year 6	**-.25**	**.10**	-.04	-.03	**-.24**	**.43**	**-.21**	.07	*-.08*	.002	**-.24**	**.32**	**-.13**	*-.09*	**.18**	*-.08*	-.02	-	
19. Year 7	**-.23**	**.11**	-.02	.01	**-.19**	**.36**	**-.22**	*.09*	-.02	.05	**-.19**	**.36**	**-.10**	**-.10**	**.12**	**-.11**	*-.08*	**.76**	-

Note. Overall $N_{Student}$ = 927 (N_{T1} = 873; N_{T2} = 831, N_{T3} = 807). T = Time; Fixed = Fixed mindset; Mastery = Mastery goal; Approach = Performance-approach goal; Avoidance = Performance-avoidance goal; Attribution = Failure-to-ability attribution; Self-con. = Self-concept; Task = Unsophisticated task differentiation; Evaluation = Public evaluation; Self-reg = Little promotion of self-regulation; Value = Differential value; Risk = Little promotion of risk-taking. Correlations take all students into account using pairwise-present. Boldfaced correlations and skewness statistics are significant at $p<.01$, italicised correlations are significant at $p<.05$.

Table 7.3

Pearson's Correlations Amongst Teachers' Motivational Beliefs

	1	2	3	4	5	6
1. Fixed mindset	-					
2. Mastery goal	-.31	-				
3. Performance-approach goal	-.12	.04	-			
4. Performance-avoidance goal	.15	.19	.15	-		
5. Failure-to-ability attribution	.35	-.26	-.07	.004	-	
6. Perceived ability/expectation	-.03	.21	.34	**.43**	-.32	-

Note. N = 30. Performance-avoidance goal had a low reliability; hence, correlations should not be interpreted. Boldfaced correlation is significant at p <.05.

Table 7.4

Pearson's Correlations Between Teachers' Motivational Beliefs and Student-Perceived FAPs

	Fixed	**Mastery**	**Approach**	**Avoidance**	**Attribution**	**P. Ability**
Student level						
Task	-.02	**.16**	**.11**	.07	.02	*-.09*
Evaluation	.04	-.02	**.15**	.02	.08	-.05
Self-reg.	**.25**	-.04	-.04	-.06	.06	.03
Value	**.18**	**-.10**	.07	*-.10*	*.09*	*-.10*
Risk	**.11**	.01	-.05	-.05	.02	*-.15*
Teacher level						
Uniform FAPs	-.20	.16	.13	-.04	-.25	.09

Note. Overall $N_{Student}$ = 808, Overall $N_{Teacher}$ = 25. Fixed = Fixed mindset; Mastery = Mastery goal; Approach = Performance-approach goal; Avoidance = Performance-avoidance goal; Attribution = Failure-to-ability attribution; P. Ability = Perceived ability/expectation; Task = Unsophisticated task differentiation; Evaluation = Public evaluation; Self-reg = Little promotion of self-regulation; Value = Differential value; Risk = Little promotion of risk-taking. Boldfaced correlations are significant at p<.01, italicised correlations are significant at p<.05.

7.1.2 Longitudinal invariance

Appendix I presents the factor loading patterns for the unconstraint configural model for motivational beliefs at Time 1 (Table I.1) and Time 3 (Table I.2) as well as the standardised factor correlations across time points (Table I.3). Table 7.5 provides the fit indices and changes in fit indices for the configural invariance, scalar invariance, and residual variance invariance models. Results indicate that all three models fit the data well. Adding increasingly restrictive constraints to the configural measurement model did not lead to significant decreases in model fit. All of the changes were well below common cut-off values proposed in the literature. It should be noted that results should be considered with caution because the proposed cut-off criteria were based on continuous data using ML or MLR estimators rather than the WLSMV estimator. Nevertheless, since the changes in fit indices were very small, and very small deviations may not necessarily affect the interpretation of results (Liu et al., 2017), it seemed reasonable to move forward. Overall, results suggest that fixed mindset, mastery goal orientation, performance-approach and -avoidance goal orientation, and self-concept were fully invariant across time at the configural, scalar, and residual variance levels. Accordingly, means of these variables could be compared across time points.

7.1.3 Model convergence

I checked multiple diagnostics for the estimation of the Bayesian regression models predicting students' motivational beliefs and achievement. All models, including those with and without custom priors, passed the checks. That is, the trace plots looked adequate, there were no divergent numerical trajectories, $\hat{R}$ values were below 1.1, there were no suspiciously high degrees of autocorrelation between samples, and the effective sample size was above 10% of the actual sample size for all parameters. Taken together, this suggests that the estimation process was adequate, and the results could be interpreted.

7.1.4 Model fit: Posterior predictive checks

Posterior predictive checks were used to assess whether the models fit the data well (for the plots for the CC models and CCMM models, see Appendix J1 and J2, respectively). The comparison between observed outcome values (y) and simulated or replicated outcome values (y_{rep}) indicated that the models fit the data appropriately although there were some simulated predicted outcome values that were slightly higher or lower than the observed outcome values.

Table 7.5

Summary of Model Fit Statistics for Testing Measurement Invariance Across Time

Model	χ^2(df)	p	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔTLI	ΔRMSEA	ΔSRMR
Configural invariance	927.68 (777)	<.001	0.988	0.984	0.016 [0.012; 0.020]	0.024	-	-	-	-
Scalar invariance	1148.99 (954)	<.001	0.984	0.983	0.016 [0.013; 0.020]	0.030	-0.004	-0.001	<0.001	0.006
Residual invariance	1221.44 (977)	<.001	0.980	0.979	0.018 [0.015; 0.021]	0.033	-0.004	-0.004	0.002	0.003

Note. CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root-mean-square error of approximation (with 90% CI); SRMR = standardised root mean residual. Δ indicates change in fit indices.

7.2 Longitudinal effects of student-perceived FAPs on students' motivational beliefs and achievement

This section presents the results for the cross-classified (CC) model with the single-teacher sample and the cross-classified multiple membership (CCMM) model with the entire sample using customised priors for the intercepts and regression coefficients and brms default priors for all other parameters. Results for models using only the default priors did not substantially differ from those with custom priors (see Appendix J.3).

7.2.1 Unconditional models

Results for the unconditional random intercept or null models are presented in Table 7.6 for the model with the single-teacher sample (CC_{NULL}) and Table 7.7 for the model with the entire sample ($CCMM_{NULL}$). In the CC_{NULL} model, the random effects (reported in standard deviations) for teachers ranged from 0.05 to 0.32 and for classrooms from 0.04 to 0.49, indicating that there was very little variation between teachers and classes for some variables (e.g., mastery goal) but a larger variation for others (e.g., achievement). For most outcomes, random effects for teachers and classrooms were very similar. However, for achievement, the random effect was larger for classrooms (0.49 [0.24, 0.77]) compared to teachers (0.19 [0.01, 0.54]). This might be because many classrooms were grouped by ability, whereas some teachers taught different ability groups. This was similar to the $CCMM_{NULL}$ model, where random effects for teachers ranged from 0.02 to 0.31 and for classrooms from 0.01 to 0.70. Random effects for teachers and classrooms were again comparable for most outcomes but not for self-concept and achievement. Interestingly, the random effect for self-concept seemed somewhat larger for teachers (0.21 [0.09, 0.32]) than classrooms (0.09 [0.003, 0.22]), whereas the random effect for achievement was larger for classrooms (0.70 [0.52, 0.93]) than teachers (0.18 [0.01, 0.50]). Hence, variability in students' self-concept was greater between teachers than between classes but variability in achievement was greater between classes than teachers.

Table 7.6

Results of the Two-Level Cross-Classified Random Intercept Model Predicting Students' Motivational Beliefs and Achievement Using the Single-Teacher Sample

	Fixed mindset	**Mastery goal**	**P.-approach**	**P.-avoidance**	**Attribution**	**Self-concept**	**Achievement**
Fixed Effects							
Intercept	1.16 [1.04, 1.28]	3.25 [3.17, 3.32]	1.34 [1.22, 1.47]	1.59 [1.48, 1.72]	-	2.70 [2.54, 2.87]	-0.001 [-0.27, 0.27]
Random Effects for Teacher (Number of levels: 17)							
Intercept	0.10 [0.01, 0.25]	0.05 [0.002, 0.14]	0.08 [0.003, 0.22]	0.07 [0.002, 0.20]	0.32 [0.03, 0.67]	0.20 [0.02, 0.41]	0.19 [0.01, 0.54]
ICC	0.02	0.01	0.01	0.01	-	0.07	0.04
Random Effects for Classroom (Number of levels: 24)							
Intercept	0.10 [0.01, 0.25]	0.04 [0.002, 0.12]	0.09 [0.004, 0.23]	0.09 [0.004, 0.24]	0.25 [0.02, 0.58]	0.16 [0.01, 0.35]	0.49 [0.24, 0.77]
ICC	0.02	0.01	0.01	0.01	-	0.04	0.25

Note. N = 342. P.-approach = Performance-approach goal; P.-avoidance = Performance-avoidance goal; Attribution = Failure-to-ability attribution. Brms default priors were used for all parameters except for the intercept. Here N(2,2) was used for all outcomes except for achievement, where N(0,1) was used. Fixed effect estimates are beta estimates and random effects are SD estimates. All estimates are presented with their 95% credible interval. Thresholds for attribution are 1) -0.94 [-1.23, -0.67], 2) -0.12 [-0.39, 0.14], 3) 0.87 [0.60, 1.15], 4) 1.62 [1.31, 1.94].

Table 7.7

Results of the Two-Level Cross-Classified Multiple Membership Random Intercept Model Predicting Students' Motivational Beliefs and Achievement Using the Entire Sample

	Fixed mindset	**Mastery goal**	**P.-approach**	**P.-avoidance**	**Attribution**	**Self-concept**	**Achievement**
			Fixed Effects				
Intercept	1.17 [1.08, 1.26]	3.23 [3.19, 3.28]	1.47 [1.37, 1.59]	1.68 [1.58, 1.79]	-	2.69 [2.58, 2.80]	-0.15 [-0.42, 0.11]
			Random Effects for Teacher (Number of levels: 29)				
Intercept	0.08 [0.004, 0.20]	0.02 [0.001, 0.05]	0.10 [0.01, 0.25]	0.09 [0.004, 0.23]	0.31 [0.15, 0.49]	0.21 [0.09, 0.32]	0.18 [0.01, 0.50]
ICC	0.01	0.001	0.03	0.03	-	0.08	0.03
			Random Effects for Classroom (Number of levels: 34)				
Intercept	0.10 [0.01, 0.21]	0.01 [0.001, 0.04]	0.17 [0.03, 0.31]	0.15 [0.02, 0.28]	0.12 [0.01, 0.32]	0.09 [0.003, 0.22]	0.70 [0.52, 0.93]
ICC	0.02	0.0004	0.01	0.01	-	0.01	0.43

Note. N = 664. P.-approach = Performance-approach goal; P.-avoidance = Performance-avoidance goal; Attribution = Failure-to-ability attribution. Brms default priors were used for all parameters except for the intercept. Here N(2,2) was used for all outcomes except for achievement where N(0,1) was used. Fixed effect estimates are beta estimates and random effects are SD estimates. All estimates are presented with their 95% credible intervals. Thresholds for attribution are 1) -0.79 [-.98, -0.61], 2) -0.002 [-0.17, 0.17], 3) 0.89 [0.70, 1.08], 4) 1.68 [1.47, 1.90].

When looking at the ICC (i.e., the proportion of variance in the outcome that is explained by clustering[18]) for both null models, the teacher and classroom clusters explained a comparable and only a small percentage of variance in fixed mindset, mastery goals, performance-approach and -avoidance goals. In the CC_{NULL} model, ICCs for these variables ranged from 1% to 2% for both teachers and classrooms. In the $CCMM_{NULL}$ model, ICCs for these variables ranged from 0.1% to 3% for teachers and 0.04% to 2% for classrooms. Accordingly, the majority of variability in these outcomes was attributable to individual student characteristics. Clusters explained slightly more variance in self-concept. In the CC_{NULL} model, the teacher cluster explained 7% and the classroom cluster explained 4% of the variation, and in the $CCMM_{NULL}$ model, this were 8% and 1%, respectively. This indicates that the teacher might be more relevant for self-concept than the class. In contrast, the classroom seemed to explain more variance in achievement than the teacher cluster. In the CC_{NULL} model, this was 4% for the teacher cluster and 25% for the classroom cluster, and in the $CCMM_{NULL}$ model, this was 3% and 43%, respectively.

7.2.2 Conditional models

The conditional CC model using the single-teacher sample is presented in Table 7.8, and the conditional CCMM model using the whole sample is presented in Table 7.9. Both models include sex (girl) and the lagged outcome variables as controls and the aggregated, uniform FAPs dimension as a predictor at the classroom/teacher level. Only the CC model has the five individual FAPs dimensions as predictors at the student level.

[18] The ICC was calculated by dividing the variance for the random effect by the total variance, that is, the variance of all random effects and the within-student (residual) variance.

Table 7.8
Results of the Two-Level Cross-Classified Random Effects Model Predicting Students' Motivational Beliefs and Achievement Using the Single-Teacher Sample

	Fixed mindset	Mastery goal	P.-approach	P.-avoidance	Attribution	Self-concept	Achievement
			Fixed Effects				
Intercept	1.19 [1.08, 1.29]	3.20 [3.10, 3.29]	1.35 [1.23, 1.49]	1.59 [1.47, 1.71]	-	2.66 [2.57, 2.75]	-0.05 [-0.19, 0.11]
Girl	0.05 [-0.18, 0.27]	**0.19 [0.004, 0.38]**	-0.01 [-0.27, 0.23]	0.11 [-0.11, 0.32]	0.25 [-0.17, 0.70]	-0.07 [-0.28, 0.11]	0.08 [-0.14, 0.28]
Lagged outcome	**0.33 [0.21, 0.46]**	**0.34 [0.21, 0.47]**	**0.44 [0.26, 0.61]**	**0.41 [0.27, 0.55]**	**0.18 [0.08, 0.29]**	**0.64 [0.53, 0.76]**	**0.69 [0.57, 0.81]**
Task	*-0.12 [-0.26, 0.03]*	-0.03 [-0.12, 0.06]	0.002 [-0.21, 0.21]	0.09 [-0.09, 0.28]	-0.03 [-0.34, 0.27]	0.02 [-0.11, 0.15]	0.01 [-0.15, 0.17]
Evaluation	0.02 [-0.18, 0.22]	-0.004 [-0.11, 0.11]	*0.12 [-0.05, 0.29]*	-0.03 [-0.27, 0.20]	0.05 [-0.31, 0.40]	0.08 [-0.07, 0.24]	-0.09 [-0.25, 0.07]
Little self-reg.	-0.02 [-0.18, 0.15]	-0.07 [-0.17, 0.03]	-0.07 [-0.26, 0.12]	0.01 [-0.17, 0.19]	-0.14 [-0.50, 0.20]	-0.03 [-0.14, 0.08]	0.02 [-0.11, 0.16]
Value	**0.19 [0.02, 0.36]**	-0.002 [-0.12, 0.12]	0.05 [-0.16, 0.24]	0.05 [-0.16, 0.25]	*0.30 [-0.06, 0.62]*	**-0.15 [-0.29, -0.02]**	0.02 [-0.16, 0.20]
Little risk	*0.15 [-0.07, 0.40]*	*-0.13 [-0.26, 0.01]*	**0.22 [0.03, 0.43]**	0.11 [-0.10, 0.35]	0.18 [-0.24, 0.61]	-0.02 [-0.16, 0.12]	*-0.13 [-0.29, 0.04]*
Aggregated FAPs	0.25 [-0.21, 0.68]	-0.31 [-0.70, 0.08]	0.21 [-0.40, 0.74]	-0.03 [-0.56, 0.53]	0.17 [-0.87, 1.25]	0.08 [-0.33, 0.49]	0.17 [-0.47, 0.78]
			Random Effects for Teacher (Number of levels: 17)				
Intercept	0.08 [0.003, 0.21]	0.10 [0.01, 0.23]	0.10 [0.004, 0.27]	0.09 [0.004, 0.25]	0.29 [0.01, 0.70]	0.06 [0.003, 0.18]	0.18 [0.01, 0.40]
Girl	0.17 [0.01, 0.47]	0.18 [0.01, 0.45]	0.19 [0.01, 0.55]	0.13 [0.004, 0.40]	0.37 [0.02, 0.98]	0.19 [0.01, 0.47]	0.14 [0.004, 0.41]
Lagged outcome	0.10 [0.004, 0.27]	0.13 [0.01, 0.31]	0.19 [0.01, 0.41]	0.13 [0.01, 0.33]	0.12 [0.01, 0.39]	0.09 [0.004, 0.24]	0.09 [0.003, 0.26]
Task	0.10 [0.004, 0.28]	0.05 [0.002, 0.14]	0.18 [0.01, 0.50]	0.14 [0.01, 0.39]	0.25 [0.01, 0.67]	0.09 [0.004, 0.26]	0.13 [0.01, 0.36]
Evaluation	0.19 [0.01, 0.44]	0.07 [0.003, 0.20]	0.09 [0.003, 0.28]	0.20 [0.01, 0.52]	0.27 [0.01, 0.74]	0.11 [0.004, 0.31]	0.11 [0.01, 0.31]
Little self-reg.	0.13 [0.01, 0.33]	0.08 [0.003, 0.21]	0.17 [0.01, 0.43]	0.12 [0.01, 0.35]	0.32 [0.01, 0.79]	0.06 [0.002, 0.18]	0.10 [0.004, 0.26]
Value	0.11 [0.004, 0.31]	0.10 [0.01, 0.25]	0.12 [0.01, 0.34]	0.13 [0.01, 0.37]	0.28 [0.01, 0.74]	0.09 [0.004, 0.24]	0.16 [0.01, 0.40]
Little risk	0.24 [0.02, 0.57]	0.10 [0.01, 0.26]	0.12 [0.01, 0.36]	0.16 [0.01, 0.43]	0.36 [0.02, 0.92]	0.07 [0.003, 0.21]	0.11 [0.004, 0.32]
			Random Effects for Classroom (Number of levels: 24)				
Intercept	0.07 [0.004, 0.20]	0.08 [0.004, 0.19]	0.12 [0.004, 0.28]	0.11 [0.01, 0.28]	0.29 [0.02, 0.67]	0.07 [0.002, 0.19]	0.16 [0.01, 0.36]
Girl	0.20 [0.01, 0.48]	0.14 [0.01, 0.37]	0.16 [0.01, 0.47]	0.13 [0.004, 0.38]	0.37 [0.02, 0.93]	0.15 [0.01, 0.39]	0.14 [0.01, 0.41]
Lagged outcome	0.08 [0.003, 0.23]	0.09 [0.004, 0.24]	0.14 [0.01, 0.35]	0.11 [0.01, 0.27]	0.18 [0.01, 0.48]	0.08 [0.003, 0.21]	0.11 [0.01, 0.26]
Task	0.09 [0.004, 0.24]	0.06 [0.002, 0.16]	0.14 [0.005, 0.40]	0.12 [0.004, 0.34]	0.26 [0.01, 0.67]	0.09 [0.003, 0.25]	0.11 [0.004, 0.30]
Evaluation	0.14 [0.01, 0.37]	0.07 [0.003, 0.19]	0.09 [0.003, 0.26]	0.15 [0.01, 0.42]	0.31 [0.02, 0.74]	0.11 [0.003, 0.31]	0.12 [0.01, 0.33]
Little self-reg.	0.17 [0.01, 0.39]	0.07 [0.003, 0.19]	0.13 [0.01, 0.36]	0.11 [0.004, 0.31]	0.35 [0.02, 0.80]	0.06 [0.003, 0.18]	0.10 [0.01, 0.25]
Value	0.10 [0.004, 0.28]	0.07 [0.003, 0.21]	0.15 [0.01, 0.38]	0.11 [0.004, 0.33]	0.19 [0.01, 0.54]	0.08 [0.004, 0.22]	0.12 [0.01, 0.33]
Little risk	0.20 [0.01, 0.52]	0.09 [0.01, 0.24]	0.12 [0.01, 0.35]	0.14 [0.01, 0.40]	0.42 [0.02, 1.05]	0.08 [0.003, 0.23]	0.10 [0.004, 0.27]

Note. N = 326. P.-approach = Performance-approach goal. P.-avoidance = Performance-avoidance goal, Attribution = Failure-to-ability attribution, Task = Unsophisticated task differentiation, Evaluation = Public evaluation, Little self-reg = promotion of little self-regulation, Value = Differential value, Little risk = Little promotion of risk-taking. Brms default priors were used for all parameters except for fixed effects. For intercepts, N(2,2) was used for all response variables except for achievement where N(0,1) was used. For regression coefficients, N(0,2) was used for all response variables except for achievement and failure-to-ability attribution where N(0,1) was used. Fixed effect estimates are (beta) estimates and random effects are standard deviation estimates. All estimates are presented with their 95% credible intervals. Thresholds for attribution are 1) -0.82 [-1.15, -0.48], 2) 0.19 [-0.13, 0.52], 3) 1.43 [1.09, 1.78], 4) 2.41 [2.02, 2.85].

Table 7.9

Results of the Two-Level Cross-Classified Multiple Membership Random Effects Model Predicting Students' Motivational Beliefs and Achievement Using the Entire Sample

	Fixed mindset	**Mastery goal**	**P.-approach**	**P.-avoidance**	**Attribution**	**Self-concept**	**Achievement**
Fixed Effects							
Intercept	1.17 [1.10, 1.25]	3.18 [3.13, 3.24]	1.44 [1.34, 1.53]	1.65 [1.55, 1.75]	-	2.68 [2.61, 2.75]	-0.12 [-0.27, 0.03]
Girl	0.05 [-0.07, 0.18]	0.07 [-0.03, 0.17]	0.05 [-0.13, 0.24]	0.06 [-0.11, 0.24]	0.25 [-0.03, 0.52]	-0.09 [-0.22, 0.04]	0.07 [-0.08, 0.20]
Lagged outcome	**0.37 [0.26, 0.48]**	**0.35 [0.27, 0.44]**	**0.45 [0.34, 0.55]**	**0.44 [0.34, 0.53]**	**0.16 [0.10, 0.23]**	**0.61 [0.52, 0.71]**	**0.66 [0.55, 0.78]**
Aggregated FAPs	0.05 [-0.25, 0.32]	-0.01 [-0.19, 0.18]	**0.52 [0.15, 0.90]**	0.27 [-0.10, 0.67]	0.39 [-0.13, 0.93]	0.01 [-0.26, 0.28]	0.11 [-0.43, 0.63]
Random Effects for Teacher (Number of levels: 29)							
Intercept	0.06 [0.003, 0.17]	0.05 [0.002, 0.13]	0.07 [0.003, 0.20]	0.09 [0.004, 0.24]	0.16 [0.01, 0.37]	0.06 [0.003, 0.17]	0.11 [0.01, 0.30]
Girl	0.09 [0.004, 0.27]	0.11 [0.01, 0.25]	0.24 [0.02, 0.55]	0.16 [0.01, 0.44]	0.33 [0.02, 0.79]	0.13 [0.01, 0.32]	0.12 [0.01, 0.31]
Lagged outcome	0.12 [0.01, 0.28]	0.08 [0.004, 0.19]	0.11 [0.01, 0.26]	0.10 [0.01, 0.25]	0.10 [0.01, 0.27]	0.08 [0.004, 0.21]	0.15 [0.01, 0.32]
Random Effects for Classroom (Number of levels: 34)							
Intercept	0.06 [0.003, 0.16]	0.04 [0.001, 0.10]	0.08 [0.003, 0.21]	0.09 [0.01, 0.24]	0.12 [0.01, 0.30]	0.07 [0.003, 0.16]	0.36 [0.23, 0.52]
Girl	0.09 [0.01, 0.25]	0.07 [0.003, 0.19]	0.16 [0.01, 0.42]	0.18 [0.01, 0.43]	0.23 [0.01, 0.58]	0.10 [0.004, 0.27]	0.12 [0.01, 0.31]
Lagged outcome	0.13 [0.01, 0.29]	0.10 [0.01, 0.20]	0.14 [0.01, 0.29]	0.08 [0.004, 0.21]	0.13 [0.01, 0.30]	0.12 [0.01, 0.24]	0.15 [0.01, 0.30]

Note. N = 625. P.-approach = Performance-approach goal. P.-avoidance = Performance-avoidance goal, Attribution = Failure-to-ability attribution. Brms default priors were used for all parameters except for fixed effects. For intercepts, N(2,2) was used for all response variables except for achievement where N(0,1) was used. For regression coefficients, N(0,2) was used for all response variables except for achievement and failure-to-ability attribution where N(0,1) was used. Fixed effect estimates are (beta) estimates and random effects are standard deviation estimates. All estimates are presented with their 95% credible intervals. Thresholds for attribution are 1) -0.48 [-0.66, -0.28], 2) 0.33 [0.14, 0.53], 3) 1.28 [1.08, 1.50], 4) 2.14 [1.90, 2.40].

7.2.2.1 Student-level effects

Considering the student-level control variables first, results show that sex predicted mastery goals in the CC model, with girls adopting more mastery goals (0.19 [0.004, 0.38]). This relationship was weaker in the CCMM model (0.07 [-0.03, 0.17]), where negative or negligible effects were not ruled out by the data. A trend in the CCMM model (0.25 [-0.03, 0.52]) suggests that girls were more likely to attribute failure to a lack of ability, but in the CC model (0.25 [-0.17, 0.70]), the uncertainty was large and negative and negligible effects were not ruled out. The results show that all lagged outcome variables positively predicted the outcome at Time 3, with regression coefficients ranging from 0.18 to 0.69 in the CC model, and from 0.16 to 0.66 in the CCMM model. Initial self-concept and achievement were particularly strong predictors of corresponding outcomes, while initial failure-to-ability attribution was a relatively weak predictor. This may indicate that the failure-to-ability attribution item was relatively unreliable, perhaps because this was a single item.

Next, I am turning to the effects of students' perceptions of the five FAPs dimensions, which were only estimated in the CC model. Besides the numerical estimates in Table 7.8, density plots for the regression coefficients are shown for all student-level FAPs dimensions predicting fixed mindset (Figure 7.1), mastery goal (Figure 7.2), performance-approach goal (Figure 7.3), performance-avoidance goal (Figure 7.4), failure-to-ability attribution (Figure 7.5), self-concept (Figure 7.6), and achievement (Figure 7.7). In these plots, the 95% CIs are shown as shaded areas under the curve, and the mean is presented as a vertical line.

Fixed mindset was positively predicted by differential value (0.19 [0.02, 0.36]). Moreover, trends suggest that little promotion of risk-taking was positively related to fixed mindset (0.15 [-0.07, 0.40]), but due to the large uncertainty around the estimate, negative or negligible effects were not completely ruled out. Given that the size of the mean coefficient was approximately half of the size for the effect of initial fixed mindset, this trend seems considerable. Unexpectedly, a trend indicated a negative relationship between unsophisticated task differentiation and fixed mindset (-0.12 [-0.26, 0.03]), but positive effects or negligible effects were not completely ruled out.

Figure 7.1
Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Fixed Mindset

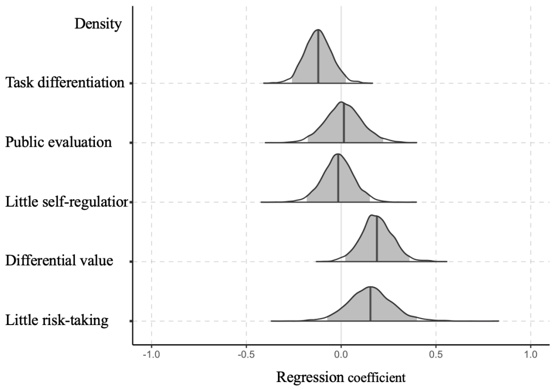


Figure 7.2
Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Mastery Goals

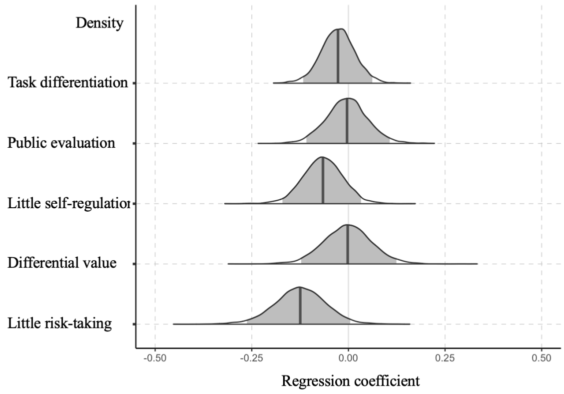


Figure 7.3
Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Performance-Approach Goals

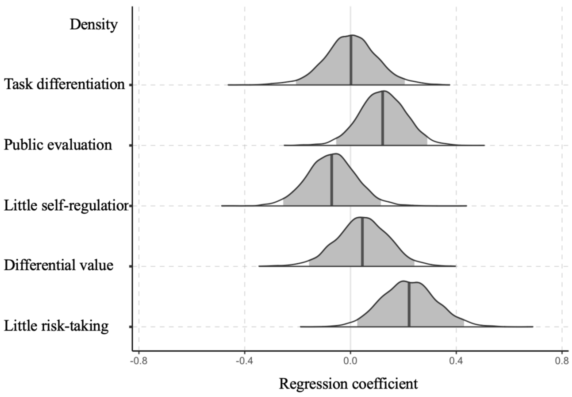


Figure 7.4
Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Performance-Avoidance Goals

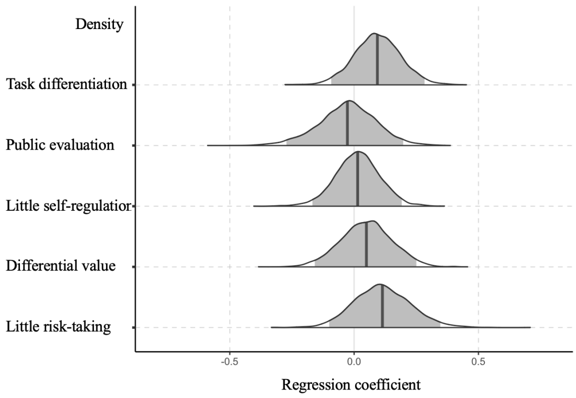


Figure 7.5
Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Failure-to-Ability Attribution

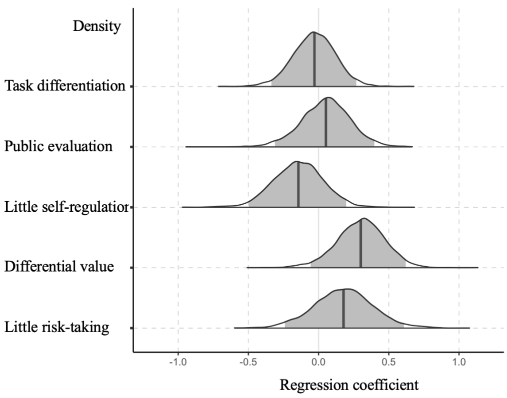


Figure 7.6
Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Self-Concept

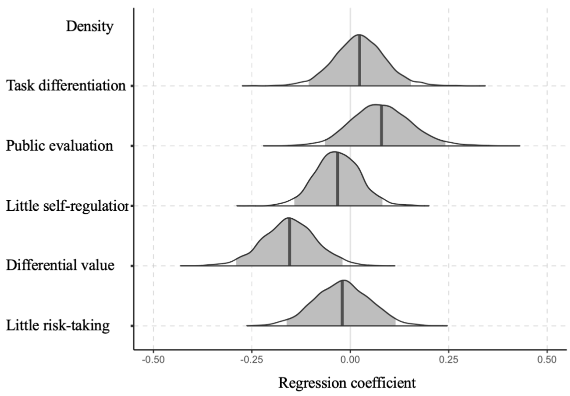
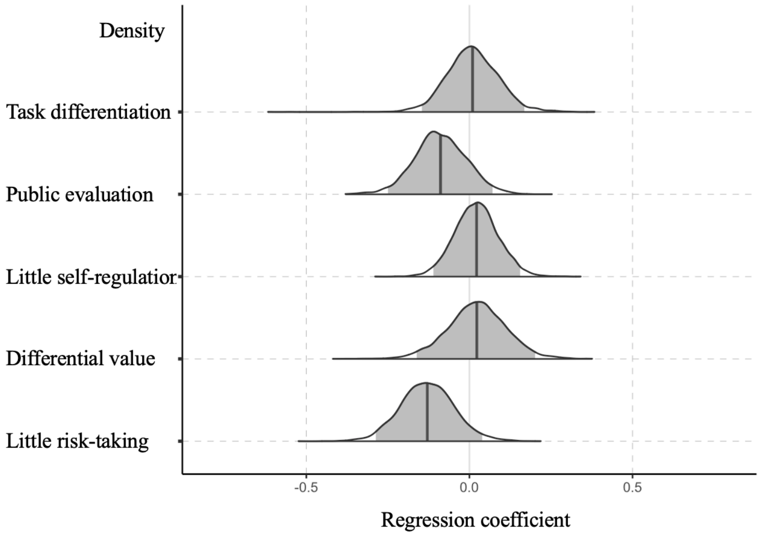


Figure 7.7

Density Plot for Regression Coefficients of Student-Level Fixed-Ability Practices Predicting Achievement

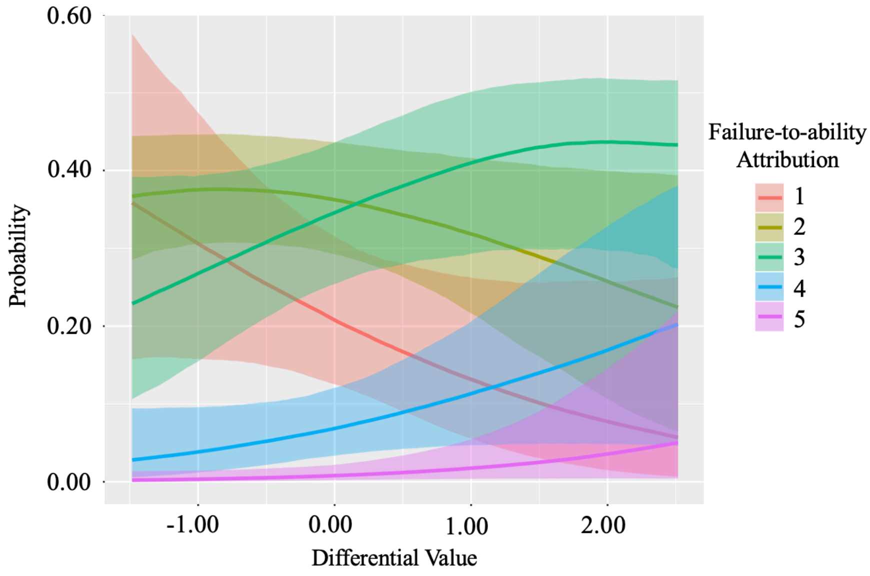


No definite evidence was found that mastery goals were predicted by any of the FAPs dimensions, but a trend suggests that little promotion of risk-taking was negatively related to mastery goals (-0.13 [-0.26, 0.01]). Yet, negligible effects were not completely ruled out. Performance-approach goals were predicted by little promotion of risk-taking (0.22 [0.03, 0.43]). Moreover, a trend suggests that public evaluation was positively related to performance-approach goals (0.12 [-0.05, 0.29]), but negative or negligible effects were not ruled out. There was no definite evidence or notable trend for any of the FAPs dimensions predicting performance-avoidance goals.

For failure-to-ability attribution, there were no definite effects, but a trend suggests that students' perceptions of differential value were positively related to the belief that failure is determined by a lack of ability (0.30 [-0.06, 0.62]). Again, negative or negligible effects were not ruled out by the data. As this coefficient is the result of a probit regression, it is not as easily interpreted. A better understanding can be gained when looking at the marginal effects plot, which shows the probability of belonging to each of the five response categories (from 0 = *Not at all true* to 4 = *Very true*) for the attribution item depending on the value of differential value, when all other predictors are at their mean (Figure 7.8). It can be seen that the probability of strongly agreeing (4) or agreeing (3) that failure is attributable to a lack of ability as well as belonging to the neutral response category (2) increased with increased differential value. In contrast, the probability to disagree (1) or strongly disagree (0) with failure-to-ability attribution decreased with increasing differential value. This indicates a positive relationship between differential value and failure-to-ability attribution.

Figure 7.8

Marginal Effects Plot for Differential Value Predicting Failure-to-Ability Attribution



Students' self-concept was negatively predicted by their perceptions of differential value (-0.15 [-0.29, -0.02]). There was no definite evidence for effects of the individual FAPs dimensions on students' maths achievement. Nevertheless, a trend indicates that little promotion of risk-taking was negatively related to achievement (-0.13 [-0.29, 0.04]), although positive or negligible effects were not completely excluded.

7.2.2.2 Classroom/teacher-level effects

Density plots for the effects of the aggregated, uniform FAPs dimension on all outcomes are presented in Figure 7.9 for the CC model and Figure 7.10 for the CCMM model. At the classroom/teacher level of the CC model, there was no evidence for effects of the uniform FAPs dimension on students' motivational beliefs or achievement. Notably, the uncertainties associated with the estimates at this level were rather large, which was likely due to the small sample size. The classroom/teacher level effects in the CCMM model with the entire sample were comparable to those of the CC model. However, this model with the larger number of groups showed that students' shared perceptions of FAPs positively predicted performance-approach goals (0.52 [0.15, 0.90]). Compared to the CC model, the estimates generally had less but still considerable uncertainty.

Figure 7.9

Density Plot for the Regression Coefficients of the Uniform Fixed-Ability Practices Dimension in the CC Model

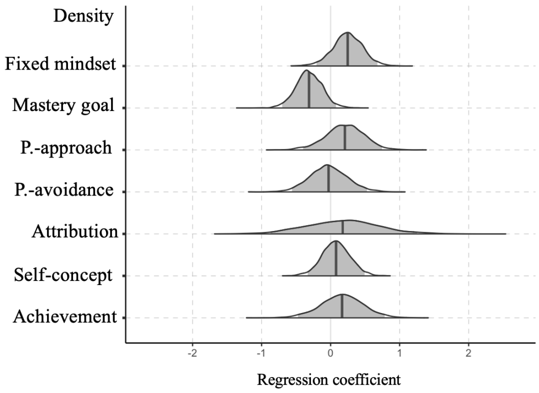
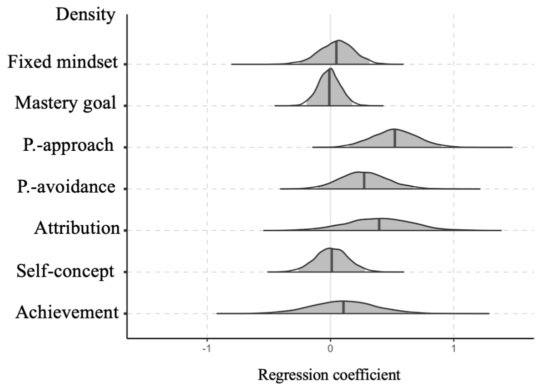


Figure 7.10

Density Plot for the Regression Coefficients of the Uniform Fixed-Ability Practices Dimension in the CCMM Model



7.3 Brief summary on RQs 2 and 3

This chapter presented the results for RQs 2 and 3. Correlations showed no relationships between teachers' motivational beliefs and students' shared perceptions of FAPs, which might have been due to a small sample size. Bayesian regression models showed that although there were few effects, the effects that were detected indicate that students' individual perceptions of more frequent FAPs fostered more maladaptive beliefs at the end of the school year while controlling for beliefs at the beginning of the school year. In particular, differential value predicted a stronger fixed mindset and a lower self-concept, and little promotion of risk-taking predicted stronger performance-approach goals. Trends suggested potential negative relationships between unsophisticated task differentiation and fixed mindsets, as well as between little promotion of risk-taking and mastery goals and achievement. Moreover, trends suggested potential positive relationships between public evaluation and performance-approach goals, between differential value and failure-to-ability attributions, and, lastly, between little promotion of risk-taking and fixed mindsets. At the teacher/classroom level, students' shared perceptions of uniform FAPs predicted stronger performance-approach goals.

Chapter 8
Qualitative Analysis Approach for Interviews
(RQs 4 and 5)

Interviews with teachers and students were analysed to explore how teachers justify their engagement in fixed-ability practices (FAPs), and whether these justifications are related to their motivational beliefs (RQ 4), as well as how students perceive teachers' FAPs, and what messages they draw from these in relation to their motivational beliefs (RQ 5). This chapter first introduces the general analysis approach using thematic analysis (Section 8.1). It then details the coding process, including how themes were developed and connections between themes were drawn (Section 8.2). The next chapter (Chapter 9) presents the identified themes and findings from teacher interviews. Findings from student interviews are presented in Chapter 10.

8.1 Overview of the thematic analysis

After considering several approaches to analysing qualitative data, such as grounded theory and interpretative phenomenological analysis, I decided to use thematic analysis (TA) to answer RQs 4 and 5. TA is regarded as a foundational approach to qualitative data analysis, and can be used to develop a rich and detailed account of the data by identifying and interpreting recurring patterns or themes (Braun & Clarke, 2006). TA has thus been described as "a form of pattern recognition within the data, where emerging themes become the categories for analysis" (Fereday & Muir-Cochrane, 2006, p. 83). A theme is defined as a specific pattern of meaning that can be identified through manifest content (i.e., explicit content at the semantic level) as well as latent content (i.e., implicit content that needs to be inferred; Braun & Clarke, 2006; Joffe, 2012). Importantly, TA can be used to describe and interpret the experiences and meaning-making processes of individuals, which is central to the research questions (Braun & Clarke, 2006).

8.2 Steps of the thematic analysis

TA is a highly flexible analysis approach, which has been argued to be a potential pitfall because it can lead to a lack of coherence. Thus, TA must be conducted in a rigorous way that lends credibility to the process and results (Nowell et al., 2017). Accordingly, I decided to perform TA following Braun and Clarke's (2006) step-by-step guide, which is one of the most influential approaches to TA (Braun et al., 2019). Their steps are widely presented as the typical process of TA and argued to enable a rigorous qualitative analysis (e.g., Neuendorf, 2019, p.213; Nowell et al., 2017; Vaismoradi et al., 2013). The steps are outlined in Table 8.1. In order to be transparent in how I analysed the data, I provided a clear audit trail throughout this chapter.

I began the data familiarisation (Step 1) with interview transcription. The audio-recordings were transcribed verbatim, including utterances and fillers. The transcripts were then imported into NVivo 12 (QSR International Pty Ltd., 2018), which was used to conduct the data analysis, and checked against the recordings. While reading the transcripts, I took notes of initial ideas and wrote a vignette for each of the twelve interviews (four teacher interviews and eight student group-interviews) to summarise the main points that were discussed (for an example, see Appendix K). Next, I began with the generation of codes (Step 2), which is now described in more detail.

Table 8.1

Steps of Thematic Analysis

Step	Description
1. Data familiarisation	Transcribing interview data; reading the transcripts several times and checking them with audiotapes; taking note of topics, first thoughts and ideas.
2. Generation of codes	Coding the entire data set systematically (mostly deductively).
3. Identification of potential themes	Identifying potential themes based on related codes.
4. Review of themes	Reviewing the themes with regard to the coded extracts and then the entire data set; generating a thematic map.
5. Definition of themes and generation of names	Further refinement of the theme and starting to analyse the story that the analysis tells so that themes are clearly defined and appropriately named.
6. Writing it up	Selection of representative and compelling quotes; final analysis of selected extracts; answering the research questions; producing the report.

Note. This table was adapted from Braun and Clarke (2006, p. 87).

8.2.1 Generation of codes

Qualitative data can be coded using an inductive approach (data-driven, bottom-up; Attride-Stirling, 2001; Patton, 2002) as well as a deductive approach (theory-driven, top-down; Boyatzis, 1998; Crabtree & Miller, 1999), and even a combination of the two, i.e., a hybrid approach (Fereday & Muir-Cochrane, 2006). I used a mainly deductive approach, and thus a theory-driven TA, given that the study was heavily based on the existing literature. Accordingly, the research questions were theory-driven, and the analysis did not aim to generate new theory. Nevertheless, I also analysed aspects that I did not previously consider if these seemed to capture something relevant to the focus on teachers' and students' experiences or perceptions of teaching practices and/or their motivational beliefs. I thus used a template of codes developed prior to the analysis based on the characterisation of the FAPs dimensions and motivational beliefs of interest, but also generated new codes based on the data itself to capture any additionally relevant aspects. The final coding framework was developed in an iterative process involving 1) an initial full round of coding, 2) the refinement of codes, 3) a second full round of coding, 4) coding part of the data by a second coder with subsequent reliability check, and finally 5) adjustment of coding as required based on reliability. These steps are described in detail in Appendix L, together with the main changes made after each step.

An example of a coded extract is presented in Table 8.2. As shown, the unit of coding was in most instances at least one speaker-turn rather than only a sentence as this helped me to contextualise the code. It was also not unusual that a code spanned multiple turns when a topic was discussed throughout (see example *moving on*). As many codes as required were assigned to a single passage. Please note that codes for teaching practices were not only applied when teachers engaged in these practices but also when students said teachers should engage in these practices (see example *mistakes for learning*). The example also shows an instance where the latent rather than manifest content was coded. I coded for *mastery goal* where the boys imply that they want to learn from mistakes and, hence, improve and "master" the learning material. Throughout the iterative coding process, I regularly checked that extracts with the same code were consistent with each other across interviewees and that the meaning of the codes did not change over time. This was mostly done by examining the individual nodes in NVivo that grouped all coded extracts together. Moreover, I grouped related codes under overarching categories and adapted these categories throughout coding. Appendix M presents the complete coding frameworks for FAPs (Table M.1) and motivational beliefs (Table M.2), including a description and an example extract for each code.

Table 8.2

Example of a Coded Extract

Extract	Codes
I: =like or what do you think about it? That she does not, like, discuss your mistake.	Moving on
B: Makes me feel not as intelligent.	Negative self-concept
B: Yeah it makes you feel dumb.	
B: 'Cause, like, it feels like that um	
B1: It feels a bit, it feels like	Mistakes for learning
B: = you haven't learned anything.	
B1: [I wish like, I wish that like]	
I: [You haven't learned anything?]	
B1: =um, she was like some other teachers, like, that, like, when someone, one person's answered, then she will explain how you've done it wrong. Because then if you don't know how you've done it wrong, you don't know how to learn from it.	Mastery goal; Discuss mistakes
I: OK. But Madam does not do that.	
B1: Madam doesn't do that, yeah.	
I: OK, and you said it doesn't f- make, sorry [someone walks in, probably the teacher and then leaves again]. So um you said it makes you feel less intelligent. Why? How?	Negative self-concept
B2: Um, like, she asks us, we answer and we get it wrong, she's like nooo, and then she goes off. Like, it's like, she's disappointed with us.	Critical of mistakes
B: Yeah.	
I: OK.	
B2: [And so and sometimes]	
B1: [Not really! No she just]=	
B2: But yeah.	
B1: =like, a tiny bit yeah but she goes like this, "No."	
B2: And then goes to someone else and then just ignores us and we are just like.	
B: Not, not disappointed but she's just telling us that it's wrong. That you are wrong.	
B2: Yeah but in maths, it like	Negative self-concept
I: [Yeah but how, yeah how does it make]	Negative emotion
B: [She doesn't say it in a mean way though]	
B2: It makes you feel a bit, like	
B: I, um, a bit down.	
B2: =intelligent, yeah.	
I: A bit down.	

8.2.1.1 Intercoder reliability

To examine the reliability of the coding process, a peer researcher coded one complete student interview and 30 minutes of one teacher interview, which is a little more than 10% of the interview data. The second coder was provided with the complete coding framework. She could also ask questions to clarify the meaning of specific codes or the coding process more generally. While it is still common to use the percentage of agreement to assess intercoder reliability (Miles et al., 2014), this is no longer recommended because it inflates reliability by not accounting for agreement occurring due to chance (see O'Connor & Joffe, 2020). Instead, I used Cohen's kappa, which takes the amount of agreement that could occur through chance into account and is as such more conservative. A Cohen's kappa between 0 and 0.20 can be considered as slight, between 0.21 and 0.40 as fair, between 0.41 and 0.60 as moderate, between 0.61 and 0.80 as substantial, and between 0.81 and 1 as nearly perfect agreement (Landis & Koch, 1977).

Cohen's kappa was calculated for the student and teacher interviews separately using a coding comparison query in NVivo. For the student interview, Cohen's kappa was 0.60, which can be considered as moderate agreement. As already noted, this statistic is a conservative measure of intercoder reliability. To compare, the percentage of agreement calculated in NVivo was above 90% for all codes. NVivo used the overlap of characters to calculate reliability, so that even when both coders agreed on a code but one included more or less context, reliability decreased. I carefully examined disagreements and discussed these with the second coder. For the most part, disagreements occurred when either I or the second coder overlooked instances rather than when we interpreted codes differently. When I had overlooked such instances, I coded these in the final round of coding (step 5 of the coding process) and also checked that I did not miss similar instances in the other interviews.

A few disagreements occurred because the second coder misunderstood the meaning of a given code. This was the case for the codes *performance-approach goal* and *performance-avoidance goal*, which the second coder applied to all instances in which students wanted to perform well as opposed to wanting to demonstrate ability or avoid demonstrating incompetence to others. In addition, *self-concept* was misunderstood as referring to more general self-esteem. Lastly, the second coder applied the code *public evaluation* to all instances in which the teacher commented on the correctness of a student's response as opposed to the teacher making a value judgement (e.g., praising the student for overall good performance or pointing out overall performance by discussing test scores). These misunderstandings were discussed and resolved.

The second coder finalised the coding of the teacher interview only after we discussed the coding of the student interview. For the teacher interview, Cohen's kappa was .54, and thus lower than the reliability for the student interview. Nevertheless, this can be considered as moderate agreement. The lowest percentage of agreement for a specific code was 89%. There were again some discrepancies regarding how much context was included during coding; I tended to code the entire speaker-turn while the second coder tended to code on a sentence-level. To compare, when using the paragraph rather than character to determine Cohen's kappa, the figure rose to .59.

When closely examining disagreements, it was again very clear that these were cases of overlooked instances on either my side or the side of the second coder. There was no particular pattern to this oversight, such as one code being frequently overlooked. No instance of disagreement due to a different interpretation of codes was detected. I carefully checked the teacher interviews for overlooked instances before proceeding with the next steps in TA.

8.2.2 Identification and definition of themes

After the final coding phase, I began with the identification of themes (Step 3). While reviewing themes (Step 4) and defining and naming themes (Step 5) are listed as separate steps in Braun and Clarke's (2006) guideline, I found that this was a much more fluid process during which I moved back and forth between these three steps. To be consistent throughout the thesis, the naming of themes was based on the theoretical framework. In particular, I used maladaptive practices to name themes relating to teaching practices, rather than adaptive ones, even if teachers predominately engaged in adaptive practices. That is, for example, even if teachers mostly encouraged help-seeking, I named the theme *little encouragement of help-seeking* or even if teachers showed little differential value, I named the theme *differential value*. As mentioned above, throughout the coding process, I grouped related codes into categories or subthemes, which can be seen in the coding framework. I adapted the categorisation until I felt that all codes found a "home" because they clearly linked to a subtheme and/or theme. This was more straightforward for some codes than others. For instance, most codes relating to the tasks that teachers assigned could be grouped relatively easily into the subthemes *unsophisticated task differentiation* and *unidimensional tasks* based on the FAPs conceptualisation. However, the code *staggered task student choice* (i.e., teachers provide tasks with different difficulty levels that students themselves can choose as opposed to the teacher deciding which students work at what level) relates to *little choice* and an adaptive way of *task differentiation*. Finally, I decided to group it with *unsophisticated task differentiation* as I felt that differentiation was more central to the code, but there are arguments for both themes.

To provide another example, the majority of codes that I ultimately grouped under the subtheme *performance over process emphasis* changed their "place" multiple times throughout the process. Most of these codes emerged during coding, and I was unsure how to group them. I initially named this subtheme *performance over process recognition*, based on the FAPs conceptualisation, but then realised that many related codes could be grouped under *performance over process emphasis*, which is more inclusive. I also grouped the code *challenge students* under this subtheme, which was not as straightforward. Challenging (or not challenging) students also relates to *little promotion of risk-taking* because it shows whether or not the teacher values struggle. After careful consideration, I thought that the purpose of challenge is to stretch students and develop mastery, which is why I grouped that code under *performance over process emphasis*.

Throughout this categorisation process, I carefully reviewed the themes and subthemes by cross-checking them with the coded extracts to ensure they were consistent and checking that I did not impose meaning based on theory that was not supported by the data themselves. The definition and naming of themes also proceeded in an iterative way that was informed by the theoretical framework.

After identifying, defining, and naming the themes, I continued to analyse the data to answer RQs 4 and 5. First, I wanted to identify connections between teachers' motivational beliefs and their engagement in FAPs and, second, I wanted to identify connections between FAPs and students' motivational beliefs. To draw these connections, I created matrices following the advice by Miles et al. (2014). Two short examples, one for teachers and one for students, are presented in Appendix N. For the teacher matrix, I plotted all statements that the four teachers made regarding a given teaching practice against each other (see Table N.1). I did not only examine how teachers' motivational beliefs related to their engagement in a given practice but considered all statements as teachers used other justifications besides their motivational beliefs. Justifications are underlined in the excerpts. Each column presents a different teacher because I was particularly interested in comparing the more-fixed and less-fixed mindset teachers. For the student matrix, I plotted FAPs against motivational beliefs and used quotes from all student interviews to show relationships between the two (see Table N.2). This helped me to identify recurring relationships or those supported by multiple students because the matrix clearly showed how many statements supported a given connection. Importantly, I also examined relationships that were not supported by many students or those that students had different opinions about because these could still reveal important information regarding the research question.

Chapter 9
Interview Themes and Findings of Teacher Interviews (RQ 4)

This chapter presents the themes relating to fixed-ability practices (FAPs) and motivational beliefs that were identified in teacher and student interviews. It also presents the findings relating to research question (RQ) 4. Specifically, teacher interviews were analysed using thematic analysis to gain greater understanding of how teachers justify their engagement in FAPs and whether their explanations relate to their motivational beliefs. The chapter begins by describing the participants and setting of the observations and interviews (Section 9.1). This is followed by a summary of the themes relating to FAPs as well as teachers' and students' motivational beliefs (Section 9.2). Subsequently, the findings for RQ 4 are presented (Section 9.3). The findings of student interviews are presented in the following chapter (Chapter 10).

9.1 The participants and setting

Before presenting the themes, it is important to briefly describe the teachers and students that participated in the observations and interviews. Four teachers participated with their classes: Ms Evans participated with a high-ability set (1 of 5), Ms Walker participated with a lower-ability set (4 of 5) although she also taught a higher set at her school, Ms Barns participated with a middle set (3 of 6), and Ms Green also participated with a middle set (4 of 6). The former two teachers taught at Meadows School and the latter two teachers taught at Woodpark School. As described in the Methodology (see Section 3.3.1.2), teachers were selected based on the motivational beliefs, in particular the mindsets, they reported in the questionnaire at the beginning of the study. Descriptive information of these teachers can be found in Table 3.4 (see Section 3.3.3.2). Throughout this chapter, I refer to teachers as "more-fixed mindset" and "less-fixed mindset" teachers. This distinction was based on teachers' mindsets in comparison to others in the sample, and differences overall were relatively small. Even the more-fixed mindset teachers did not strongly endorse a fixed mindset, they merely did not reject it as strongly as other teachers in the sample. The more-fixed mindset teachers Ms Evans and Ms Walker had fixed mindset scores of 1 and 1.75, respectively, on a scale from 0 to 4. The less-fixed mindset teachers Ms Barns and Ms Green had scores of 0 and 0.25, respectively, and thus very strongly rejected a fixed mindset. In addition, the more-fixed mindset teachers had slightly stronger performance-avoidance goals but did not differ notably in other respects from the other two teachers. Students were also selected based on their motivational beliefs at Time 1 as well as the performance levels their teachers reported at the time of interviews (see Section 3.3.1.2).

During interviews, teachers and students often pointed out that a given practice was school policy. It is thus important to bear in mind that the discussed practices were embedded within the larger school context. For instance, at Meadows School, the school of the more-fixed mindset teachers, all teachers had to follow the same lesson plan and, thus, had relatively little autonomy over their activities. In fact, teachers described that the maths department decided together what materials and slides to use, although it seemed that they could adjust tasks or materials to their ability groups. Hence, Ms Evans and Ms Walker frequently used the pronoun "we" rather than "I" when describing their practices. They also mentioned school policies such as celebrating mistakes and doing *"everything with a purpose."* Moreover, they explained that their school followed a maths mastery programme, which aimed at developing mastery rather than fluency. Ms Barns and Ms Green made comparatively few mentions of school policies. Examples included that the school aimed to foster students' resilience, and Ms Barns sometimes referred to the behaviour policy, which outlines sanctions for students' poor behaviour, such as detention. Additionally, teachers at both schools could give students merits or demerits for good or poor performance and behaviour, respectively. For instance, at Meadows School, teachers should give a demerit to students whose test performance declined, and award merits to the "fab four" students, including students that worked particularly hard during the lesson.

Lastly, it should be pointed out that some teaching practices were discussed in more detail in some interviews than others due to time constraints. Thus, there was sometimes an imbalance in how much information participants could provide about a given practice.

9.2 The themes: Fixed-ability practices and motivational beliefs

This section provides a summary of the themes that were identified in teacher and student interviews. This synopsis aims to show what practices teachers engaged in according to teacher and student interviews and what motivational beliefs teachers and students endorsed. As such, it acts as an introduction that sets the scene before addressing RQs 4 (this chapter) and 5 (next chapter). Appendix O provides a more detailed description of the themes including specific examples and quotes representing FAPs (O.1) and motivational beliefs

(O.2). The themes closely mirrored the overarching constructs in the study, which was perhaps not surprising as I focused my interview questions on FAPs, and because I conducted a mainly theory-driven thematic analysis. New important aspects that I did not consider prior to the interviews were also identified, which usually fit within one of the broader constructs of theoretical interest.

9.2.1 Fixed-ability practices

Figure 9.1 depicts the themes and subthemes relating to FAPs (for an overview of the FAPs themes and corresponding codes, see the coding framework in Appendix M, Table M.1). Six themes were identified: tasks (unsophisticated task differentiation, unidimensional tasks), grouping (group and pair work, seating and ability grouping, stability), evaluation (public evaluation, performance over process emphasis), little autonomy (little choice, little promotion of self-regulation), differential value, and little promotion of risk-taking (little encouragement of help-seeking, unconstructive handling of mistakes). The subtheme *group and pair work* was not specified a priori and the subtheme *performance over process emphasis* has been extended to encompass new aspects relating to evaluation practices. In addition, *school policy* emerged as an important contextual aspect within which the other practices were embedded. An overview of the extent to which teachers engaged in FAPs according to their own and student interviews is presented in Table 9.1.

Figure 9.1

Thematic Map of Fixed-Ability Practices

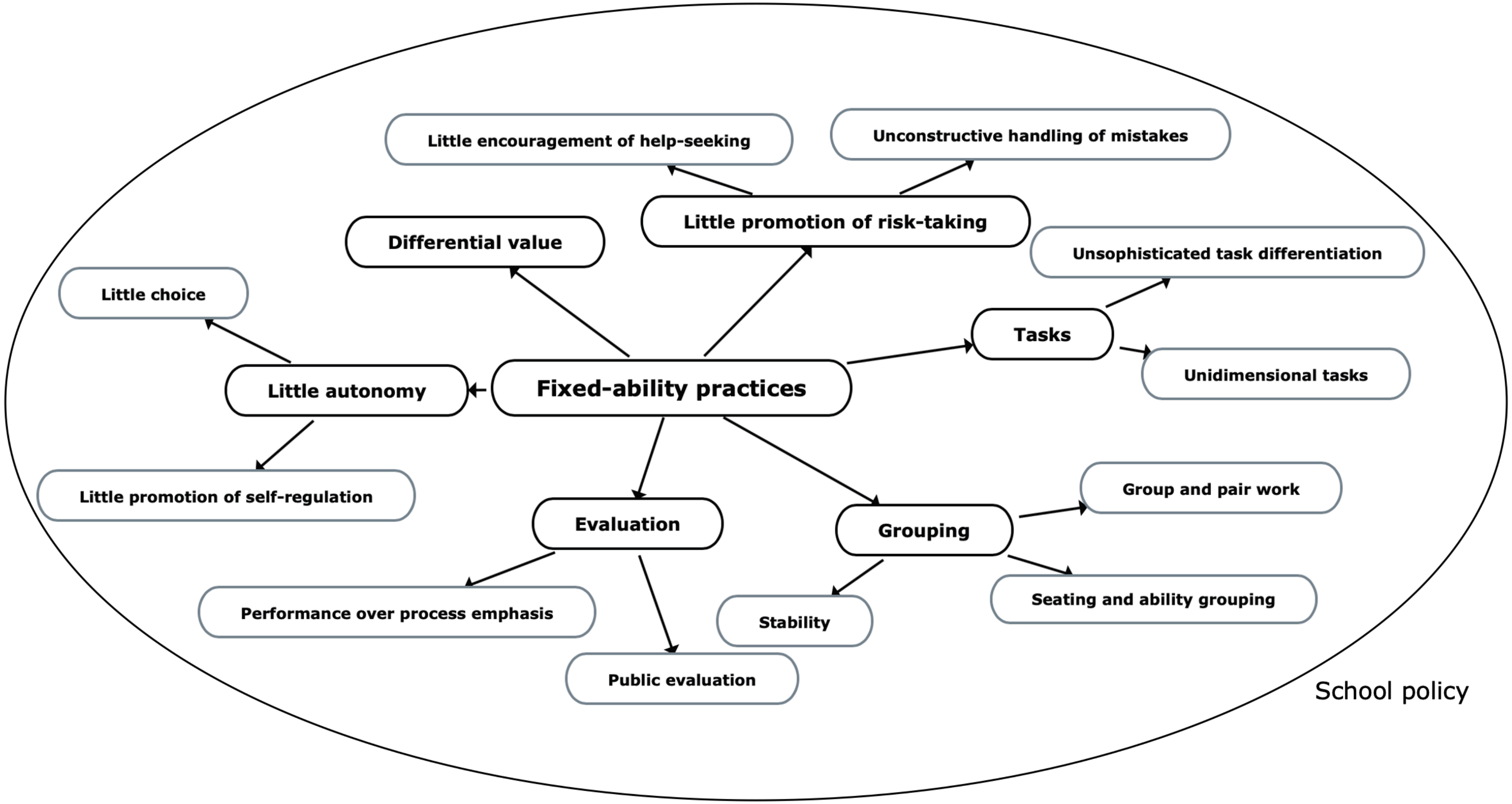


Table 9.1

Teachers' Fixed-Ability Practices Reported in Teacher and Student Interviews

Dimension	Ms Evans	Ms Walker	Ms Barns	Ms Green
Tasks				
Unsophisticated task differentiation	More growth	More growth	More growth	More growth
Unidimensional tasks	More growth	Mixed[a]	Mixed	More growth
Grouping				
Group and pair work	[b]More fixed	[b]More fixed	More growth	More growth
Seating and ability grouping (<u>in class</u>)	Mixed	Mixed	Mixed	Mixed
Stability (<u>school-wide</u>)	[c]NA	[c]NA	[c]NA	[c]NA
Evaluation				
Public evaluation	More fixed	More fixed	More growth	Mixed
Performance over process emphasis	Mixed	Mixed	Mixed	More growth
Little autonomy				
Little choice	More fixed	More fixed	Mixed	More growth
Little promotion of self-regulation	Mixed	Mixed	More growth	More growth
Differentiation value	More growth	More growth	More growth	More growth
Little promotion of risk-taking				
Little encouragement of help-seeking	[d]Mixed	[d]Mixed	[e]Mixed	More growth
Unconstructive handling of mistakes	More growth	More growth	More growth	More growth

Note. Colour coding: green = more growth; yellow = mixed; red = more fixed.
[a]Teacher reported fewer unidimensional tasks but her students reported they mostly had to follow her methods.
[b]Teachers reported more partner work but students reported, and observations showed, that students engaged in very little pair work.
[c]This was a school policy that teachers had no control over.
[d]Teachers reported that students could help each other but students reported, and observations showed, that this was rarely the case.
[e]Teacher reported that she answered questions thoroughly, but some students reported she did not engage with questions in detail and they did not feel comfortable to ask for help.

9.2.1.1 Tasks

The *tasks* theme encompasses all aspects relating to the kind of tasks that teachers assigned and how these tasks needed to be solved. It has two subthemes, namely *unsophisticated task differentiation*, which relates to the teacher assigning fewer or easier tasks to lower-performing students, and *unidimensional tasks*, which relates to aspects such as whether tasks could be solved in only one way and had only one correct solution, and whether or not teachers provided extension tasks for students who finished early. According to teachers and students, none of the four teachers seemed to engage in unsophisticated task differentiation. Only the less-fixed mindset teachers offered tasks with different difficulty levels, but they let students decide what level to work on. Yet, Ms Barns reported that she usually asked all students to begin with the easier tasks. All four teachers taught and allowed students to use different methods to solve the same tasks, although Ms Barns did not seem to actively teach different methods. In addition, while Ms Walker reported she would usually let her students choose which methods to use, and while her students said they were allowed to use their own method if they were confident enough, they also said they mostly had to follow Ms Walker's way. All teachers seemed to make sure that no student sat without work when finishing tasks early by providing them with extra work.

9.2.1.2 Grouping

The *grouping* theme encompasses the subthemes *group and pair work*, which addresses whether classes engaged in partner or group work, *seating and ability grouping*, which addresses how teachers designed the seating plan, especially whether it was based on students' performance, and whether classes were grouped by ability, as well as *stability*, which addresses whether groupings were flexible or fixed. The less-fixed mindset teachers offered students many opportunities to work and discuss in pairs, and Ms Barns even offered group activities. In contrast, the more-fixed mindset teachers offered relatively few opportunities for students to work together. Most teachers decided on the seating plan randomly. However, all but Ms Barns moved students around for behaviour management, such as seating distracted students next to well-behaved ones, seating weaker students towards the front, or pairing weaker and stronger students. Ms Barns opposed such practices. She was also the only teacher who did not show a clear preference for heterogeneous pairs. Students at both schools were set by ability but these sets were more flexible at Meadows School, where students were re-set every half term instead of only at the end of the school year.

9.2.1.3 Evaluation

The *evaluation* theme encompasses all aspects regarding how students are being evaluated and what goals the teachers emphasise. It has two subthemes, *public evaluation*, which relates to whether students are being evaluated in private or in front of the class, and *performance over process emphasis*, which relates to whether teachers emphasise aspects such as performance and speed or mastery and effort. While all teachers engaged in some kind of public evaluation, the more-fixed mindset teachers did so more frequently and more explicitly. For example, they asked students to read out test scores in front of the whole class. All teachers emphasised the importance of effort and deep understanding and tended to give process-oriented praise. The more-fixed mindset teachers seemed to emphasise the importance of performance more than the less-fixed mindset teachers, but Ms Barns also emphasised speed. In general, performance emphasis was stronger at Meadows School, where students took tests very frequently.

9.2.1.4 Little autonomy

The *little autonomy* theme encompasses all aspects relating to how much autonomy support teachers provide to their students. It is divided into the subthemes *little choice*, which regards whether students can make their own decisions during lessons, and *little promotion of self-regulation*, which considers whether students evaluate their own learning, set their own goals, and engage in self- and peer assessment. Students had, overall, less choice in classes of the more-fixed mindset teachers, where the lessons were highly structured, and tasks had to be completed in a set amount of time. The less-fixed mindset teachers provided more choice, such as allowing students to decide what difficulty level they wanted to work at and whether they wanted to work in pairs or alone. All four teachers asked students to engage in self- and peer-assessment, but only the less-fixed mindset teachers asked students to engage in self-evaluation, such as reflecting on their performance and thinking about areas they needed to improve.

9.2.1.5 Differential value

The theme *differential value* regards mostly whether all students are encouraged to participate during lessons or whether the teachers mostly engage the high-performing students. Another aspect is whether the teacher shows equal interest in all students, although this was not discussed in detail during interviews. There were no marked differences between teachers regarding showing differential value to students based on their performance. They seemed to value students equally and encouraged all to participate, especially those that were weaker or rarely volunteered answers. Yet, teachers did not always succeed in calling on all students equally and they sometimes even picked students who they expected to know the answer straight away.

9.2.1.6 Little promotion of risk-taking

The theme *little promotion of risk-taking* relates to the extent to which teachers make students feel safe to show that they are unsure about their maths work. The subtheme *little encouragement of help-seeking* regards whether teachers encourage students to seek teacher and peer support, and the subtheme *unconstructive handling of mistakes* relates to how teachers react when students make a mistake, such as engaging in discussion or quickly moving on to the next student. All teachers offered help to their students, but the less-fixed mindset teachers more actively encouraged students to help each other. Some students reported that they did not feel comfortable to ask Ms Barns for help. None of the teachers minded mistakes and instead regarded them as learning opportunities, but the more-fixed mindset teachers seemed to even celebrate mistakes. Lastly, all teachers showed some constructive handling of mistakes, such as discussing mistakes and giving students the chance to try again, and some unconstructive handling, such as immediately moving on to another student.

9.2.2 Motivational beliefs

Figure 9.2 depicts the themes and subthemes for motivational beliefs (for an overview of the motivational beliefs themes and corresponding codes, see the coding framework in Appendix M, Table M.2). Five motivational beliefs themes were identified as well as one theme relating to teachers' conception of the nature of maths and one theme relating to students' emotions. Themes and subthemes were mindset (growth mindset, fixed mindset), achievement goal orientation (mastery goal, performance-approach goal, performance-avoidance goal), attribution (ability attribution, effort attribution, situation attribution), student self-concept, teachers' perceived ability/expectation for students, nature of maths, and student emotion. Newly identified were the themes *situation attribution, nature of maths*, and *student emotion*. Overviews of teachers' and students' motivational beliefs based on the interview data are presented in Table 9.2 and Table 9.3, respectively. See Appendix O for a detailed description of these themes besides the summary provided in this chapter.

Figure 9.2

Thematic Map of Motivational Beliefs

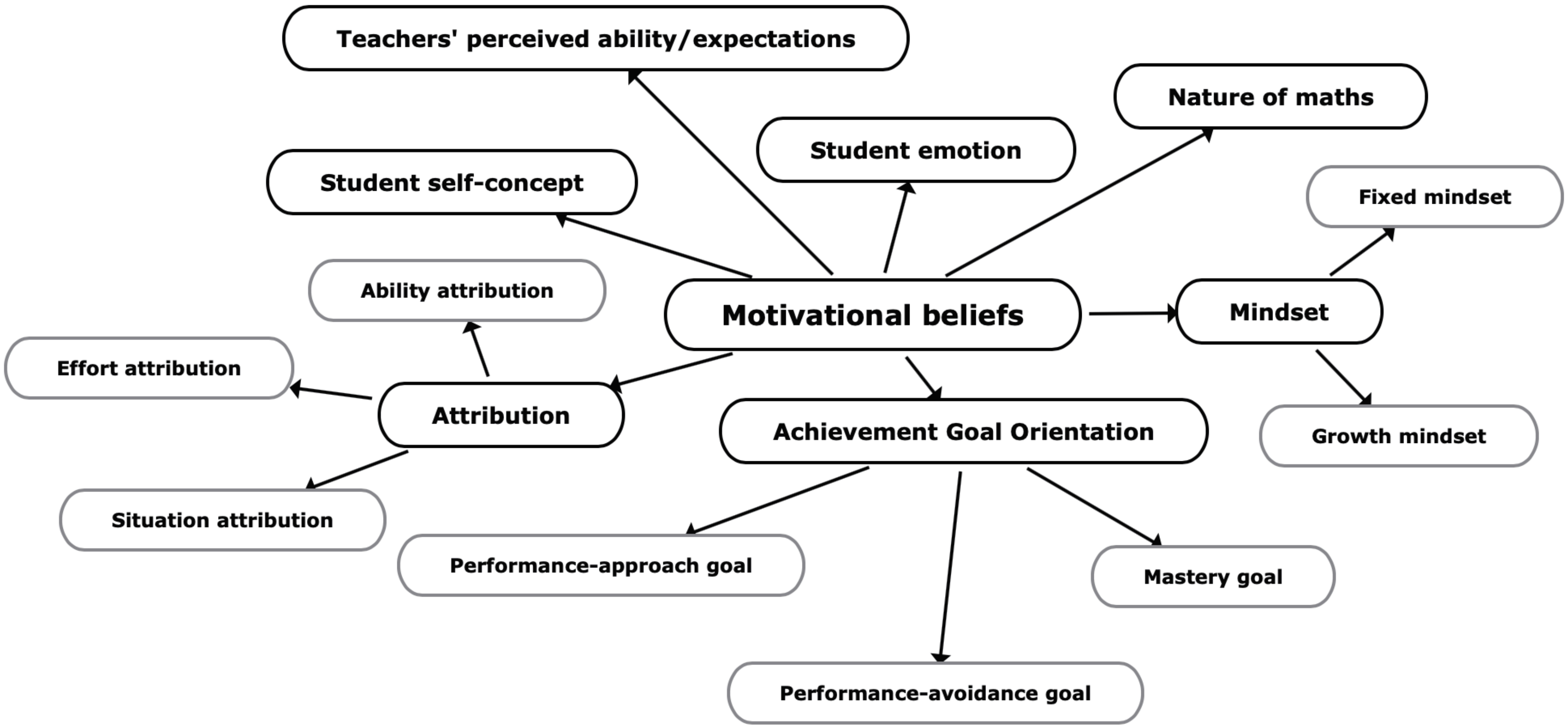


Table 9.2

Teachers' Motivational Beliefs as Reported in Interviews

Motivational belief	Ms Evans	Ms Walker	Ms Barns	Ms Green
Mindset				
Growth	Some	Some	Some	Some
Fixed	NA	Some	Some	Some
Goal orientation				
Mastery	NA	NA	NA	NA
Performance-approach	NA	NA	NA	NA
Performance-avoidance	NA	NA	NA	Few
Attribution				
Effort	Some	Some	Some	Some
Ability	NA	Some	Some	Some
Situation	NA	Few	NA	Few
Perceived ability/expectation	Mostly high	Mixed	Mostly high	Mixed
Nature of maths	Complex	Mostly procedural	Complex	Complex

Note. NA = not appropriate because beliefs were not revealed. Colour coding: green = adaptive, yellow = neutral, red = maladaptive.

Table 9.3

Students' Motivational Beliefs as Reported in Interviews

Motivational belief	Ms Evans' class	Ms Walker's class	Ms Barns' class	Ms Green's class
Mindset				
Growth	Some	Some	Some	Some
Fixed	NA	NA	NA	NA
Goal orientation				
Mastery	Strong	Strong	Some	Strong
Performance-approach	Strong	NA	NA	Few
Performance-avoidance	Strong	Some	NA	NA
Attribution				
Effort	Yes	Some	Some	Some
Ability	Some	Some	Some	Some
Situation	Some	NA	NA	Few
Self-concept	Mixed	Mixed	Mixed	Mixed
Emotion	Mixed	Mixed	Mixed	Mixed

Note. NA = not appropriate because beliefs were not revealed. Colour coding: green = adaptive, yellow = neutral, red = maladaptive.

9.2.2.1 Mindset

The *mindset* theme, with its subthemes *growth mindset* and *fixed mindset*, encompasses all instances in which individuals showed whether they believed that improvement was possible, even if they referred to improving performance and understanding rather than ability.

9.2.2.1.1 Teacher mindset

All four teachers predominantly showed beliefs in a malleable ability. They indicated that students could improve with effort despite having different abilities. Nevertheless, all teachers but Ms Evans, one of the more-fixed mindset teachers according to her questionnaire responses, also exposed some fixed-mindset beliefs. For instance, Ms Walker did not seem to believe that many of her students in the lower-ability set could move upwards as she thought they were not in that set because of their behaviour. Similarly, Ms Barns thought that most students stayed on the same level or in the same *"pecking order"* during their school career. Ms Green also stated that she did not believe her students in a lower-ability set would take A-level maths, although these students were only in Year 7.

9.2.2.1.2 Student mindset

Students only exhibited growth mindset beliefs. They thought, for instance, that they could improve if they worked hard, revised, and got challenged. Two girls from Ms Walker's class actually reported they learned during a school assembly that they were supposed to have a growth mindset.

9.2.2.2 Achievement goal orientation

The *achievement goal orientation* theme has the three subthemes *mastery goal* (e.g., wanting to learn and improve), *performance-approach goal* (e.g., wanting to demonstrate good performance or high ability), and *performance-avoidance goal* (e.g., not wanting to show poor performance or low ability).

9.2.2.2.1 Teacher goal orientation

Ms Green was the only teacher who revealed some goal orientations during interviews. She expressed some performance-avoidance goals, but these instances were very rare. This, of course, does not mean that this was her only goal orientation or that the other teachers endorsed no goal orientation. These were merely not expressed during interviews.

9.2.2.2.2 Student goal orientation

All students expressed mastery and performance goals in similar ways but the extent to which goals were endorsed differed between classes. All students endorsed mastery goals, but students in Ms Evans' class also expressed performance goals particularly frequently. Mastery goals were expressed as wanting to learn, wanting a challenge, and wanting to improve. Performance-approach goals were mostly expressed as wanting to appear smart and showing off, such as showing how many correct answers one got. Performance-avoidance goals were expressed as not wanting one's classmates or teacher to know that one struggled, which was often accompanied by cheating behaviour, such as faking the number of correct answers. Notably, many students did not express any performance goals but only very few explicitly said they did not care about their performance appearance.

9.2.2.3 Attribution

The *attribution* theme regards how teachers and students explain students' performance. Three subthemes were identified, namely *ability attribution*, *effort attribution*, and *situation attribution*.

9.2.2.3.1 Teacher attribution

Teachers used a combination of ability and effort attributions to explain students' performance. They recognised that students have different abilities, which could explain their performance, but they also believed that they could improve if they worked hard. Some also offered situational explanations for poor performance.

9.2.2.3.2 Student attribution

Similar to teachers, most students made both ability and effort attributions. They asserted that students had different abilities and learned at a different pace but that investing effort and working hard could help students to improve. There were also some instances of situation attribution, such as having too much homework and too long school hours, but these were scarce.

9.2.2.4 Student self-concept

The theme *student self-concept* regards how students view their ability or performance in maths. I did not ask students directly what they thought about how good at maths they were, but students often commented on this when asked how certain FAPs made them feel.

Independent of FAPs, only a few instances occurred in which students commented on how confident they were about their maths work. Girls tended to indicate more often than boys that they were struggling, while boys were rather confident about their performance.

9.2.2.5 Teachers' perceived ability and expectations for students

This theme encompasses how teachers perceive their students' ability and what expectations they have for them. Overall, all teachers described their students as having different abilities and as performing differently in their classes. Ms Evans, who taught the highest set, and Ms Barns, who taught a middle set, seemed to have the highest expectations. It was surprising that it was mostly one of the less-fixed mindset teachers, Ms Green, who seemed to have had lower expectations for her students overall and who also communicated differential expectations more explicitly. This might be related to her teaching a lower set. Yet, Ms Walker, who taught a similar set, did not show such strong tendencies although she also did not think that many of her students could move upwards.

9.2.2.6 Nature of maths

The theme *nature of maths* emerged from the data and was not considered a priori. All teachers revealed during interviews what they thought maths was like without being prompted to do so. They described how everything in maths was connected to something that came before, and that maths was not about memorising but understanding. Ms Barns also emphasised that maths should be fun and interesting to do. Although there were not many instances of teachers discussing these aspects, it seemed relevant for teachers' practices (see Section 9.3).

9.2.2.7 Student emotion

The theme *student emotion* also emerged from the data and was not considered a priori. When students were asked what they thought about certain teaching practices, they often responded by describing their emotional response. Students expressed a range of positive and negative emotions and also made judgements about the fairness of given practices. Sometimes, they also reported being indifferent towards a given practice.

9.3 Teachers' justifications for engaging in fixed-ability practices

This section presents the results for RQ 4, that is, it explores how teachers justify their engagement in FAPs and whether these justifications are related to their motivational beliefs. It is again important to note that not all practices were discussed with all teachers to the same extent due to time constraints, so that some teachers have not provided justifications for some of the practices. An initial overview of the justifications is provided in Table 9.4.

9.3.1 Teachers' motivational beliefs and justifications for tasks

9.3.1.1 Unsophisticated task differentiation

None of the four teachers engaged in unsophisticated task differentiation. Teachers' justifications for this differed, but there were no systematic differences in explanations between more- and less-fixed mindset teachers. Task differentiation was not discussed during interviews with Ms Walker. Hence, only the justifications of the other three teachers were examined. Ms Evans and Ms Barns did not assign easier tasks to some students because they did not want to set up different expectations. Ms Evans stated that she used task differentiation in the past but would no longer do this. She explained:

> *I think that it is very difficult, it also sets up your expectations before they've even tried. So, if you say, like, you are gonna do A and you're gonna do B, you are saying, well, you are never gonna be able to do that one, so I'm just gonna give you this task.*

Now she wanted to have *"high expectations for everybody."* Ms Barns sometimes provided tasks with different difficulty levels and let students choose what level to work at. She explained she did not want to determine who should work at what task because *"the pupil who suddenly joins the dots (...) can go, can stream ahead."* Similar to Ms Evans, this indicates that she did not want to set up expectations before students started because they might suddenly improve. It also points towards this practice being determined by her belief in growth.

Table 9.4

Teachers' Justification for Their Engagement in Fixed-Ability Practices Revealed During Interviews

Dimension	Ms Evans	Ms Walker	Ms Barns	Ms Green
Unsophisticated task differentiation	**No:** → sets up expectations → other types of differentiation → similar abilities	**No:** → no justification provided	**No, but give choice:** → sets up expectations → growth mindset → other types of differentiation	**No, but give choice:** → different abilities → value/foster autonomy
Unidimensional tasks	**No:** → foster mastery **Sometimes:** → performance concerns	**No:** → foster mastery → student preferences → reduce struggle	**Sometimes:** → high expectations	**No:** → student preferences
Group and pair work	**Sometimes:** → foster mastery **No:** → low expectations → performance concerns → school policy	**Sometimes:** → foster mastery **No:** → low expectations → learners as recipients of knowledge → performance concerns → school policy → need structure	**Yes:** → foster mastery	**Yes:** → foster mastery
Seating and ability grouping (in-class)	**Heterogeneous:** → foster mastery → foster growth **Weak at front:** → lower expectations	**Heterogeneous:** → foster mastery → foster growth **Weak at front:** → lower expectations	**Heterogeneous:** → foster mastery → foster growth **Homogenous:** → foster mastery	**Heterogeneous:** → foster mastery → foster growth **Keep homogenous pairs:** → foster mastery
Stability (school-wide)	NA	NA	NA	NA
Public evaluation	**Yes:** → practical considerations → positive error climate → healthy competition → formative assessment	**Yes:** → motivate students → interfere with student perceptions **Private criticism:** → protect students' feelings	**Sometimes:** → practical considerations → formative assessment **No:** → protect students' feelings	**Sometimes:** → motivate students

Performance over process emphasis	**Process:** → value/foster mastery → value/foster effort → different abilities → school policy **Little performance praise:** → avoid performance-avoidance goals	**Process:** → value/foster mastery → value/foster effort → different abilities → effort attribution → motivate students → school policy	**Process:** → value/foster mastery → value/foster effort → different abilities → nature of maths → motivate students	**Process:** → value/foster mastery → value/foster effort → nature of maths → motivate students → school policy
Choice	**Little:** → importance of structure → performance concerns → foster mastery → teacher expertise	**Little:** → importance of structure → performance concerns → foster mastery	**Yes:** → value autonomy	**Yes:** → value/foster autonomy
Promotion of self-regulation	**Sometimes:** → no justification provided	**Sometimes:** → foster mastery	**Yes:** → understand students' opinions → school policy	**Yes:** → foster mastery → value/foster autonomy → reduce teacher workload → school policy
Differential value	**No:** → foster mastery **Sometimes:** → create pace	**No:** → foster mastery	**No:** → foster mastery **Sometimes:** → questions too difficult for some	**No:** → foster mastery **Sometimes:** → questions too difficult for some
Encouragement of help-seeking	**Yes:** → no justification provided	**Yes:** → foster mastery	**Yes:** → foster mastery	**Yes:** → foster mastery
Handling of mistakes	**Constructive:** → positive attitudes towards mistakes → foster mastery → school policy **Sometimes unconstructive:** → create pace	**Constructive:** → positive attitudes towards mistakes → foster mastery → school policy **Sometimes unconstructive:** → effort attribution	**Constructive:** → positive attitudes towards mistakes **Sometimes unconstructive:** → no justification provided	**Constructive:** → positive attitudes towards mistakes → foster mastery → situation attributions **Sometimes unconstructive:** → no justification provided

Note. Colour coding: green = adaptive, yellow = neutral, red = maladaptive.

Another similarity between Ms Evans and Ms Barns was that both thought task differentiation was not necessary because they differentiated in other ways. Ms Evans explained, *"we differentiate here through time taken and through support."* Ms Barns, one of the less-fixed mindset teachers, argued that task differentiation was not necessary *"because the differentiations come in before they are doing the written work."* She further described, *"when I start, I target anyone. And it slowly comes down to, to the stronger kids."* This suggests that the "weaker" students might be left behind and could not follow everything that Ms Barns explained, which she seemed to accept. In contrast, Ms Evans seemed to differentiate by support while students worked on tasks, which might help all students to progress in their learning. Both teachers thought that all students should at least be able to access the tasks, while the stronger students could move on to more difficult parts. Thereby, they already pointed towards the use of multidimensional tasks, which is addressed in the next subsection.

Ms Evans and Ms Green both referred to ability setting to justify their practices, but they had a rather different understanding of sets. Ms Evans argued that task differentiation was not needed because students were grouped by ability, indicating she assumed most students in the same set would work on a similar level and be able to solve similar tasks. In contrast, Ms Green indicated that she believed students performed at different levels despite being in the same set. She often provided tasks with different difficulty levels that students could choose themselves, explaining, *"I want to differentiate the teaching. This is Set 4, but still some pupil[s] are closer to Set 3, even Set 2. Some are closer to Set 5 and 6."*

Lastly, Ms Green justified her decision that students could choose what difficulty level to work on with the importance of student autonomy. She explained, *"I want them to take the ownership of their learning rather than let teacher* [sic] *tell them everything (. ...) I like to train them to be independent learner[s]."*

9.3.1.2 Unidimensional tasks

All four teachers taught and allowed students to use different methods to solve the same task, but Ms Barns, one of the less-fixed mindset teachers, seemed to do this to a lesser extent. Both Ms Evans and Ms Walker provided multidimensional tasks to foster students' mastery. Ms Evans sometimes used tasks in which students could choose the numbers they used, justifying, *"we try to develop mastery, not fluency."* She also stressed that knowing different methods was important for students' conceptual understanding. She explained that knowing how you can arrive at the same solution using different methods was *"developing their, kind of their map of maths."* Ms Walker also taught different methods so that students could gain a better conceptual understanding and, hence, master the material. For example, she argued, *"in order for them to understand fractions and decimals and percentages, they need to themselves be aware that they are essentially, like, all [a] proportional thing, like, they're all the same thing."* Ms Evans revealed that performance concerns would be a reason not to teach multiple methods. She emphasised that she would only teach methods that worked in all contexts because otherwise, *"it's not a very good method because they might get confused when they come to an exam question that doesn't work."* She also admitted she would not take the time to explore different methods with year groups that were close to important exams.

Ms Walker and Ms Green both offered different methods because they thought students could then use the method they preferred. Ms Walker further explained, *"I just think they need to realise that there are different options"* and students can use *"whichever [method] they sort of immediately lean towards."* Her reason for this was that she did not *"want them to be finding it difficult if they don't have to."* Ms Barns revealed quite a different idea of what maths should be like. She repeatedly argued that maths was supposed to be difficult and challenging for students. She said students came from primary school knowing certain methods, such as column multiplication, and wondered*, "why would I dumb that down? Um, division, the same thing. Long division, why on earth would I do them the disservice?"* Thus, in contrast to Ms Walker, Ms Barns did not want to make it easier for students, so that they would not take a step backwards. Nevertheless, she accepted different methods as long as these were correct and also explained that the class usually came up with a method together.

9.3.2 Teachers' motivational beliefs and justifications for grouping

9.3.2.1 Group and pair work

All teachers engaged their students in some pair work, but the less-fixed mindset teachers did so to a greater extent than the more-fixed mindset teachers, and only Ms Barns offered any group work. All teachers explained that students could develop greater understanding through partner discussions and learn from each other, indicating that teachers offered pair work to foster mastery. For instance, Ms Evans explained she used pair work because *"talking about your answers can develop your, your reasoning skills."* Similarly, Ms Walker stated that she would use partner work so that students learned from each other. She thought students who explained something to another gained greater understanding through the action of explaining, while the students who received the explanation could *"catch up and learn."* Both less-fixed mindset teachers provided very frequent pair work because they thought it benefited students' learning because they could help each other. In addition, Ms Barns emphasised that pair work was *"nice because they can talk. But they don't realise I have set the agenda of what they are talking about."* Similarly, she tended to provide group work as a *"fun activity,"* which she thought students deserved because they worked very hard.

Interviews with the more-fixed mindset teachers revealed that low expectations and performance concerns might inhibit teachers from offering pair work. Ms Evans said she always did one talk task during lessons with the highest ability set because she thought *"they don't need to bring everything down on paper."* This suggests that she might not have done this with lower sets, assuming that they would not remember what they had learned without writing everything down. Similarly, Ms Walker assumed that only the more able students could teach anything to other students as she said, *"I don't think there's any point of getting two children who don't understand to discuss because it's a complete waste of their time."* She also explained, *"the chance that a Year 7 student is gonna discover how to do something in maths is very unlikely and I just think you could waste a whole lesson."* She argued that if you would offer group work, you should at least provide students with all the necessary knowledge, indicating she regarded learners as recipients rather than creators of knowledge. Moreover, Ms Walker stated that she did not use pair work towards the end of half term when students had to revise for tests, indicating she valued test performance over mastery.

Neither Ms Evans nor Miss Walker used group work. Ms Evans explained:

> *(...) one of our teaching and learning non-negotiable policies here is "EWAP," which is every[thing] with a purpose, and um, straightest line. So, what is the most direct route from where you are to the objective. And group work tends to not be the straightest line.*

Ms Walker also added that because the school followed the maths mastery programme, which focused on conceptual understanding, group work was less important. She thought that if the school would not follow that programme, group work would be more important because students gained a *"more real understanding of what they are doing 'cause they have to think."* Ms Evans also argued that in groups, *"it's more difficult for everyone to get heard… so it's kind of an accountability measure and also because, um, and also a behaviour management tool. So, it's not something in this school that we particularly use."* These two teachers also thought that group work did not provide enough structure for students.

9.3.2.2 Seating and ability grouping

All teachers seated students in random pairs, but all except Ms Barns changed pairs over time. All but Ms Barns paired students for behaviour management, such as seating a distracting or "chatty" student next to a well-behaved one, so that students concentrated better.

All teachers supported seating people of different abilities together with the justification that students *"can help each other."* This indicates that valuing mastery and believing that students can improve may lead teachers to form heterogeneous pairs. Ms Barns and Ms Green mentioned that they would not separate students with similar performance if they worked well together, *"[could] pull each other along,"* and were *"challenging each other."* In contrast, as already mentioned, Ms Walker did not believe that two students who were both not performing well could help each other. Moreover, the more-fixed mindset teachers seated *"weaker"* students at the front of the class to check on them more regularly and to make sure they understood everything, which suggests that they had different expectations for their students.

Ability grouping at the school level was not in the control of teachers. Hence, none of the teachers provided justifications for such ability grouping.

9.3.2.3 Stability

How often students could change ability sets was determined by the departments and not individual teachers. Hence, teachers did not discuss reasons for the stability of ability groups. Yet, all but Ms Barns changed their seating plans as explained above. Reasons for such changes mostly included managing behaviour and improving performance. For instance, Ms Walker sometimes changed pairs after examining students' test performance in the hope that they would perform better in a different pair. Ms Barns strongly opposed seating for behaviour management, saying, *"that's just not fair."* She argued that ensuring students were well-behaved was *"the teacher's job! It's not a pupil's job."*

9.3.3 Teachers' motivational beliefs and justifications for evaluation

9.3.3.1 Public evaluation

The more-fixed mindset teachers seemed to engage in more frequent and more explicit public evaluation than the less-fixed mindset teachers. Reasons for engaging in public evaluations differed between teachers and included, amongst others, practical reasons, conducting formative assessment, wanting to set good examples to get other students working, and even influencing how students perceived each other.

Beginning with practical reasons, Ms Evans explained that she asked students to read out their test performance in front of the class because checking their scores in private *"takes so much more time."* When I asked whether this created a sense of competition in the class, she responded, *"yeah but they also get, they, they know, they know it's a competitive environment because they know that they are set, that they are re-set every half term based on how they do."* Moreover, she thought that because mistakes were not regarded as something negative in the school, students would know there was *"nothing wrong with not being, not being the top. Um, but equally there's nothing wrong with healthy competition either and wanting to, to push on."* She further described that at Meadows School, they *"like to publicly show them that, you know, that there is a bit of competition."* Thus, while she engaged in public evaluation for practical reasons, she did not mind that this also created competition amongst students. Ms Barns also had practical reasons for engaging in public evaluation. Sometimes, she wrote students' names on the board to later award them a merit, explaining she otherwise *"won't remember."* In addition, she said that she was more open about demeriting than meriting because she gave demerits immediately when misbehaviours occurred, whereas she would usually *"sit down and think, 'well now, who tried hard?'"* after the lesson to award merits.

Both Ms Evans and Ms Barns engaged in some public evaluation to gain an understanding of how students were currently performing. Ms Evans said she checked how many correct answers students had on a given task, explaining, *"it's just a general, like, a really quick snapshot of, like, am I going too fast, am I going too slow… Just a couple of times a lesson just to literately check."* Similarly, Ms Barns sometimes checked and commented on students' work on mini-whiteboards in front of the class, explaining, *"I want to know how much they've learned. I want to know what they are up to. So some, some kids are slower at writing ... so yeah, getting, so I'm getting in, this is the formative feedback."*

Ms Walker and Ms Green both explained they used public evaluation and social comparison so that other students got motivated to work. Ms Walker almost constantly commented on students' performance and behaviour, which she described as:

> *(…) narrating the positives … it's just a design to get everyone on task basically. Because if four people are sat there and one of them is not behaving, you don't wanna constantly be saying 'you need to do this, you need to do this.' But if you say to people*

> *around them 'well done, you, you knew,' then that person is suddenly the one doing the right thing.*

She thought that if she would point out students who were not doing well, *"they'll just switch off and it's not like a nice, pleasant environment."* Moreover, she gave merits for good behaviour to students who generally had behavioural difficulties during the lessons *"because they need, like, those kids usually need really fast rewards."* She further explained she gave merits publicly because *"more of them will participate if they know that they have a chance of getting a merit at the end."* She also often compared the class' performance with the performance of her higher set, explaining *"that makes them firstly realise that they are pretty much on the same lessons as everyone else. And also, they just see it as a challenge. They think that it's like a competition."* Ms Green also revealed that she praised students in public to motivate other students to work. She said:

> *(...) they don't see each other's progress. They don't have comparison and some people tend to be lazy. And also, when I say it, they notice the teacher is observing them. I'm watching them. So it's kind of um telling them, 'you better work harder 'cause I'm watching you!' Lots of children are, they need, ah, they need teachers ... kind of push from time to time.*

In general, Ms Walker stood out as a teacher who used praise and merits very strategically to influence students' behaviour and to influence students' perceptions of each other. She publicly praised students, especially weaker ones, explaining, "*it's about, like, tapping into the perceptions of each other and stuff... it's a chance for me to be, like, 'don't be thinking [student] is not really clever because she is.'*" In line with this, she gave criticism in private because she "*wouldn't embarrass them in front of everyone*." In contrast, Ms Barns stood out as a teacher who strongly opposed comparing students or discussing the performance of individual students. She declared she would *"absolutely not"* share students' test performance because that was "*cruel*." She justified this with her personal experience, explaining, "*once I got the lowest mark and the teacher read back the marks…and she gave me such a bollocking in front of everyone. I would never do that."* She also opposed students themselves sharing their performance in front of the class because *"that's just bragging. It's not, it's not pleasant. And ... what about the child behind you who's struggling? Um, no."*

9.3.3.2 Performance over process emphasis

Teachers tended to focus on effort, behaviour, and improvement rather than performance when they praised students, which was mostly due to teachers valuing these aspects and thinking that students had more control over this than actual performance, which might depend more on ability. Teachers also used effort praise as a motivational tool. In addition, conceiving maths to be about deep understanding rather than following rules seemed to explain their emphasis on mastery.

All four teachers showed that they valued student mastery. As already mentioned, Meadows School followed the maths mastery programme. In line with this, both Ms Evans and Ms Walker explained it was important to focus on mathematical reasoning and conceptual understanding and not just fluency. However, it was difficult to distinguish these teachers' opinions from the school policy because they often referred to "we" instead of "I" when talking about mastery emphasis. Ms Barns also thought it was not enough for students to follow rules; she wanted to ensure that students understood what they were doing because *"maths isn't a recipe."* Hence, her conception of the nature of maths seemed to influence what she valued in students. Likewise, she explained*, "we've got to challenge them but it, it's part of maths."* She also thought, *"otherwise [maths] is boring… you will see how dead everyone looks."* Ms Green also reported she aimed to develop students' reasoning skills, which was why she often let them explore content rather than to give them the rules in advance. She explained:

> *(…) they got used to be told everything straight away but I just thought I want to develop their thinking skills 'cause memorise things, yes quick and, look, make the lesson look um structured and also um looks like as if they are all understood it* [sic]. *But they could forget so easily.*

Teachers' mastery emphasis also became apparent in their expectation that students should always explain how they got their answers instead of only stating the results, both verbally and in writing, so that they could check if students understood. If students did not show their workings, Ms Barns explained, *"no matter that you're good [at maths], you can't do well at maths."* She also remarked that they otherwise threw away points in their tests, which Ms Green mentioned as well, and that checking their workings was a way to ensure that they were not just copying answers. In addition, Ms Green described, *"maths although it's numbers but it's also a way of communication. How you present yourself to help people who don't, don't understand."* Hence, her understanding of the nature of maths seemed to influence her desire that students explained their answers. She also thought proper *"working outs"* would help students during revisions.

All four teachers recognised students for their effort through praise and merits because they thought investing effort and being resilient was an important skill. At Meadows School, students even received effort grades. But even besides that policy, Ms Evans wanted to recognise students' effort. She explained:

> *I've taught kids before where their attainment in their test is always really poor … but you try and you, and then I think that's a valuable…the fact that you have learned and put in effort will carry you far.*

She preferred speaking about resilience rather than effort because *"it's really hard when you are getting it wrong every time, to pick yourself up… that's such an important skill to be able to do that."* Similarly, Ms Walker regarded resilience as a very important skill that deserved praise. Ms Barns also focused on effort and progress praise, explaining, *"I expect them to work hard. So I think it's important that I recognise that they have."* However, in contrast to the policy at Meadows School, Ms Barns did not *"see much point in giving an effort grade. Because we are teaching them to think so we've got to be able to track what they know."* She further said:

> *I think it's really worthwhile to acknowledge the effort, of course! I always would. But I do think we have to say, well this is where you are at. (...) And we've got a very funny system in Britain where we always tell them they are doing well, they're always succeeding.*

Ms Green also thought resilience was a very important life skill. She also explained that being resilient was part of the school rules.

Ms Walker and Ms Barns revealed that they praised effort rather than performance because students had different abilities. Ms Walker explained, *"I'm rewarding progress not just you're getting all the answers right (. …) because they need to realise that everyone is a slightly different ability. And it's, you will get there if you put in the hard work."* She also said, *"it is really hard to sit there for 50 minutes working through something that is quite difficult if you are not a particularly high attainting student."* Accordingly, her justification for effort praise lies in her belief that effort can lead to improvement and, thus, also her belief in growth. In addition, she thought *"the students that have found it harder and who have got there deserve to be praised more than the ones that have found it really easy straight away."*

Similarly, Ms Barns explained:

> *I would default mainly to effort because kids have different ability so it's quite hard if, er, well, I try not to do it but it's quite hard to only comment those who actually have got it. When actually someone else is working jolly hard and just maybe can't.*

Likewise, Ms Evans did not agree with the school policy of giving demerits to students when their performance declined, explaining:

> *I don't particularly like doing it mostly because it can often punish [the] ones you don't want to punish. Like what you want is the ones you know are being lazy. But often, you feel really bad because you know they are actually just working really hard and sometimes the numbers are just a bit trickier.*

Teachers also used effort praise as a motivational tool. Ms Walker wanted to encourage more effort through her praise because she believed that performance in lower-attaining classes was *"linked [to] how much actual effort they've put."* This indicates a relationship to effort attribution. In addition, she wanted to praise students who were *"introverted [because] they can go unnoticed in school."* Ms Barns also used effort praise as a motivational tool. She described:

> *(…) maths is hard because we are teaching them to think (. …) if we are not legitimately saying you are doing a good job, you are making progress, it becomes quite a hard task. But if you've been told that you are doing well, um, chances are you are more motivated to carry on.*

For her, genuineness was crucial, and she emphasised that she would only praise students if they really deserved it. Ms Green also used effort praise to show students that they had done well and to encourage good behaviour. She explained, *"I know it was a very hard lesson, hard for them to sit there, and um, for so long and, um, I want to encourage those positive points. Emphasise what they have done well."* Because Ms Evans taught a class with students who *"are really good anyways. They are quite motivated,"* she thought, *"they don't really, yeah, need a lot of praise."*

In general, Ms Walker praised very frequently, explaining, *"then they think, 'oh, when I do the right thing, I get praised for it.' Not spend five days thinking, 'I'm doing the right thing why is no one praising me?'"* In contrast, Ms Evans explained, *"you don't want to always praise. Like, well you've got the answer right, well done. Because then they feel bad if they get the answer wrong."* While she did not say so explicitly, it seemed as though she did not want to foster performance-avoidance goals. Ms Green also did not frequently praise students, which she justified with her cultural background, describing, *"in China, we think particularly for young children, they need to have that, um, ability to take that criticism ... rather than taking lots of praise, can't bear with any criticism."*

9.3.4 Teachers' motivational beliefs and justifications for little autonomy

9.3.4.1 Little choice

The more-fixed mindset teachers gave students fewer opportunities for choice than the less-fixed mindset teachers. The less-fixed mindset teachers appeared to value students' autonomy and being independent learners, while the more-fixed mindset teachers regarded structure and clear instructions as more important because they thought these made learning more efficient.

Ms Evans, one of the more-fixed mindset teachers, believed that students learned most efficiently when they knew exactly what the teachers expected of them. Hence, she regarded providing a lot of structure as very important. She thought that when having a clear structure, students can think and learn without distractions, and they also "*learn to work with a bit of pace because they know they are gonna have seven minutes on that task and that is that… they actually get a decent amount of work done."* This suggests that Ms Evans valued performance and speed over depth of understanding. She also argued that following a clear structure ensured that teachers kept to a high standard and did not relax rules to make lessons easier for them. In addition, she emphasised that teachers in her school followed the latest research and were mostly educated to a Master's level, which gave them the expertise to decide how students should be learning. This indicates that Ms Evans valued teacher authority over student autonomy.

Ms Walker also said it was very important to provide structure, especially for students with special educational needs, because she thought that many students did not have enough structure at home. Moreover, she argued that students enjoyed structure because *"it creates a calm environment where, like you said, they feel very, um, safe and comfortable 'cause they know what's, wha-, they know what the deal is."* While she explained that structure provided an environment in which students could learn well and do maths, she also explained, *"if you can do it relatively quickly then you're gonna enjoy and the lesson makes you feel clever kind of thing."* She thereby moved the focus from learning to performance, speed, and feeling clever.

Ms Barns and Ms Green provided more choice to students than the more-fixed mindset teachers. Ms Barns did not provide many justifications for this, but she indicated that she valued students as independent thinkers. Hence, she wanted to give them choices and let them figure out their own strategies or let them discover themselves what the topic was, explaining, *"I didn't want them to plug into that programming. I wanted them to think."* Similarly, Ms Green emphasised repeatedly that she considered it important that students were independent learners. She also highlighted that she wanted to guide students rather than tell them exactly how to do their work. She explained, *"I like to train them to be independent learner[s] but they have got used to be[ing] fed … learning. That conflicts you can see."* She also explained, *"I want them to take the ownership of their learning rather than let [the] teacher tell them everything."*

9.3.4.2 Little promotion of self-regulation

All teachers asked students to engage in self- or peer-assessment of tasks, but only the less-fixed mindset teachers asked students to engage in self-evaluation, such as reflecting on how they were doing and what they needed to improve. Neither Ms Evans nor Ms Barns offered justifications for self- and peer-assessment but the others focused on benefits to students' learning. Ms Walker thought that it helped students to *"find out where [they've] gone wrong."* She also explained that it benefitted students to provide feedback to each other and that such feedback was more immediate compared to having to wait until the end of the week when she could check their books. This indicates that she aimed to improve students' understanding through self-assessment. Similarly, Ms Green explained, *"they will look at the question in more details. If we mark it, they just look at the score,"* and *"that way they are actually learning more about from* [sic] *the test rather than just look[ing] at the score only."* This shows that she wanted students to focus on mastery rather than performance. Ms Green also explained that it reduced the teachers' workload and helped students to take more responsibility, which again highlights that she valued student autonomy. Additionally, it was a policy at Woodpark School to ask students to engage in self-assessment at the end of each lesson, although students reported that not all teachers did this as frequently as Ms Barns and Ms Green.

Interestingly, when I asked Ms Walker whether students engaged in self-evaluation or set their own goals, she said this was not the case and added, *"but I do think there are certain kids who know that I'm watching out for certain things in them."* Hence, she immediately brought attention to her setting the goals for students rather than students setting goals for themselves, which indicates that she did not value students' own goals as much. Ms Barns asked students to engage in self-evaluation, explaining, *"I want some feedback from them so that when I take their books in and mark them, actually I've got a, an idea of what they think (. ...) so that I can understand them better."* In addition, she told students, *"I want to know what you are thinking,"* indicating that she valued students' opinions and found it important to understand how they perceived themselves. Ms Green did not justify engagement in self-evaluation, but this practice is in line with her general goal to help students become independent learners.

9.3.5 Teachers' motivational beliefs and justifications for differential value

All teachers tried to encourage all of their students, especially the lower-performing or shy ones, to participate during lessons. This dimension also encompasses whether teachers showed equal interest in all students, but this was not discussed with teachers themselves. There were no marked differences between teachers regarding how they justified asking mostly the "weaker" students to give answers in front of the class. They did so because they wanted that all students were engaged and could learn and, hence, because they valued student mastery. Ms Evans explained that this was also a matter of *"accountability. Like, if they know that you're never gonna pick them, they're not gonna think of the answer (. ...) if I want them to be learning, they need to be thinking."* Similarly, Ms Walker said this ensured that students were listening, which was also in line with Ms Barns' justification for calling especially on students who were not participating. She argued, *"they should be involved in what's going on in class and I need to know why aren't you taking part. And also just to keep that overall spread."* In addition, Ms Green explained if *"you can see they are distracted or not paying attention you pick on them and to get their attention back."*

Teachers sometimes called more frequently on higher-performing students to move the lesson on more quickly. Ms Evans described, *"if I want to be quick about it then I might ask a question ... saying, like, oh I know you are going to get it right."* Importantly, she recognised *"[t]hat's not good teaching practice (. ...) Because you shouldn't really be asking questions just to get the answer (. ...) You should be asking questions to get them to think."* The less-fixed mindset teachers said that as the questions got harder, fewer and fewer people were able to answer them, so that the weaker students could often not participate anymore. Both asserted they would still try to engage them.

9.3.6 Teachers' motivational beliefs and justifications for little promotion of risk-taking

9.3.6.1 Little encouragement of help-seeking

All teachers encouraged students to ask them for help. However, the less-fixed mindset teachers seemed to encourage students to help each other more frequently than the more-fixed mindset teachers. Ms Evans did not provide justifications for encouraging help-seeking. The other teachers made clear that they wanted to ensure that all students understood what they were teaching, so that they could do the work. Accordingly, these teachers encouraged help-seeking because they valued student mastery. Ms Walker wanted students to ask her questions because otherwise, she explained, *"I'm almost myself trying to predict what they don't understand. And I think it's more useful for them to just tell me what they don't understand."* She also justified this by asking, *"[h]ow are they gonna get all they want out of the lesson or they need out of the lesson if they can't ask me, um, to help them?"* Ms Barns encouraged students to ask for help to ascertain that they understood what she taught because otherwise *"you cannot set a class off working."* Both Ms Walker and Ms Barns stated that they would not be impressed if students needed help or asked questions because they were not listening. Ms Green thought it was important that students asked her for help, although she preferred that they did so individually rather than while she taught at the board because they would otherwise hold back the rest of the class. The teachers encouraged peer support because they believed that students could learn from each other. In addition, Ms Green explained that students would have to wait too long for support if they could only ask her because she was *"the only teacher."*

9.3.6.2 Unconstructive handling of mistakes

All teachers showed a range of adaptive reactions to mistakes, such as discussing them or giving students time to try again, as well as maladaptive reactions, such as moving on to the next student or providing the correct answers themselves. It seemed as though the positive reactions were more prevalent because all students thought their teachers did not mind them making mistakes. Teachers seemed to constructively handle mistakes because they did not regard mistakes as something negative. Instead, they saw the positives in mistakes and thought they were important for learning. Similarly, teachers thought mistakes were a normal part of the learning process because all people made mistakes. Ms Walker noted that many students *"will be really worried about making mistakes… They think it's embarrassing. They think it's bad. They think it means they're stupid,"* and explained that she wanted to interfere with these maladaptive views. Another reason for Ms Evans' and Ms Walker's positive reactions to mistakes was that it was school policy to *"celebrate mistakes loads."* Ms Walker thought that this showed students that *"it's much better to just be honest."* These teachers also discussed mistakes so that students *"hopefully won't ever make it because we've discussed why it's not correct."* Hence, they focused on student mastery. Similarly, Ms Green thought making mistakes was important because *"otherwise you could only understand it superficially."* Ms Green justified her lack of criticism for mistakes with situation attributions for mistakes. She assumed she made the tasks too difficult and explained, *"there is something I haven't explained… I just jumped."*

Teachers admitted that they did not always take the time to discuss mistakes because, as Ms Evans described, *"in any lesson you are balancing about 100 different things in your brain at once. And sometimes you just need to move on."* In addition, Ms Walker was critical of *"lazy"* mistakes or mistakes that occurred due to a lack of effort rather than a lack of understanding.

9.3.7 Summary on teachers' justifications for their teaching practices

Overall, discussions with teachers revealed that their motivational beliefs indeed explained the teaching practices they engaged in, and that adaptive practices were mostly justified by adaptive beliefs. Teachers did not refer explicitly to their mindset, and they did not refer to their own achievement goals but rather to the aspects they valued in their students. Importantly, beliefs did not always seem to consistently translate to teaching practices. Besides beliefs, teachers used many other justifications such as school policies and practical considerations.

None of the teachers used unsophisticated task differentiation, and there were no marked differences regarding the justifications that teachers with more- or less-fixed mindsets provided. Justifications for not using unsophisticated task differentiation included not wanting to set up different expectations for students, having a growth mindset, using different types of differentiation instead, and students having similar abilities because they were set. The less-fixed mindset teachers provided tasks with graded difficulty levels but let students choose what level to work at. Reasons included that students had different abilities despite the sets and that teachers valued student autonomy. All four teachers taught and allowed students to use different methods to solve the same task, but Ms Barns, one of the less-fixed mindset teachers, did not teach different methods as frequently. Reasons for using multidimensional tasks included wanting to foster students' mastery, allowing students to use the methods they preferred so that they did not need to struggle, and believing that students could exceed teachers' expectations. In contrast, performance concerns might deter teachers from using different methods as indicated by one of the more-fixed mindset teachers.

All teachers engaged their students in some pair work, but the less-fixed mindset teachers did so more frequently, and only Ms Barns used any group work. Teachers thought that offering partner or group work would develop students' reasoning skills and understanding and, hence, mastery. The more-fixed mindset teachers revealed several explanations for not offering group or pair work, including low expectations, viewing learners as recipients of knowledge, performance concerns, wanting to provide more structure, and school policy. All teachers supported heterogenous ability pairs because these would foster mastery and growth. The less-fixed mindset teachers also supported, or at least did not oppose, homogenous pairs if these aided students' mastery. The more-fixed mindset teachers tended to seat weaker students towards the front of the class, which seemed to be based on the expectation that these students needed frequent support and, thus, having lower expectations for some students. The stability of school-wide ability groups was not discussed because this was out of teachers' hands. Within classes, all but Ms Barns changed the seating plans. Reasons included improving students behaviour and/or performance.

The more-fixed mindset teachers engaged in more frequent and more explicit public evaluation. Reasons for engaging in public evaluation varied, but no systematic differences were found between more- and less-fixed mindset teachers. Justifications included motivating students to work, practical reasons, and making formative assessments. Not engaging in public evaluation was justified by wanting to protect students' feelings, such as avoiding embarrassment. All teachers tended to emphasise the process over performance, for example by providing process-oriented praise. Reasons included that they valued these aspects and seemed to think that students had more control over these aspects than actual performance because students differed in their abilities. Teachers also used effort praise as a motivational tool. Conceiving maths not to be about following procedures but rather about conceptual understanding also explained an emphasis on mastery, although one of the more-fixed mindset teachers did not seem to follow this consistently.

The more-fixed mindset teachers provided less choice, which they justified by the importance of structure and clear instructions for efficient and fast learning, hence, placing performance and speed over depth of understanding. One also argued that teachers could make better decisions for students because of their expertise. The less-fixed mindset teachers justified providing choice by valuing student autonomy. All teachers asked students to engage in self-and peer-assessment to improve their understanding and, thus, mastery. Only less-fixed mindset teachers asked students to evaluate their learning. This was school policy, and Ms Barns also explained that she did this to understand how students perceived themselves.

None of the teachers showed much differential value. Instead, they tried to call on all students, in particular those who did not participate as much, because they wanted to ensure that all students were engaged and learned and, thus, to foster their mastery. One of the more-fixed mindset teachers justified calling higher-performing students by having to move the lesson on more quickly, and the less-fixed mindset teachers, surprisingly, said that not all students could always participate because questions sometimes got too difficult.

All teachers encouraged help-seeking because they wanted to ensure that all students could understand the lesson. Hence, they wanted to foster mastery. Teachers reported mostly constructive handling of mistakes, which they justified by positive attitudes towards mistakes, wanting to foster mastery, making situational attributions for mistakes, as well as by school policies. More- and less-fixed mindset teachers were also similar in that they did not always handle mistakes constructively, but only the more-fixed mindset teachers provided explanations for this. One said she did not always discuss mistakes to create pace, and the other explained that she reacted negatively to mistakes if these were due to a lack of effort rather than understanding.

Chapter 10
Findings of Student Interviews (RQ 5)

This chapter presents the findings for research question (RQ) 5, which is the final RQ of this study. Student interviews were analysed to gain insights into how students perceive fixed-ability practices (FAPs) and, in particular, what messages they draw from these for their motivational beliefs. The findings focus on relationships between FAPs and motivational beliefs that were revealed by multiple students, although I sometimes point out when an individual student made a link that seemed relevant. There were also multiple instances where students did not agree with each other or felt indifferent towards a given practice, which was also noted. An initial overview of the relationships that were revealed is provided in Table 10.1.

10.1 Students' perceptions of fixed-ability practices in relation to their motivational beliefs

10.1.1 Tasks and students' motivational beliefs

10.1.1.1 Unsophisticated task differentiation

Discussions with students revealed relationships between teachers' unsophisticated task differentiation and students' mastery goals and self-concept. A consistent response was that it would be unfair if some students got easier tasks because these students would have fewer opportunities to learn. Students regarded getting easy tasks and not being challenged as pointless. Comments such as *"I wanna do harder work 'cause then you are not really learning much at all"* and if *"you keep giving them things that they find easy, then I don't think they would, um, improve"* were very common. While it was not said explicitly, this points towards a potential relationship between unsophisticated task differentiation and a decrease in mastery goals. Notably, there were a few students who regarded this kind of differentiation as fair because *"some people could really struggle more than others."* One girl also remarked that the easy tasks might contain all the information students needed to then work on harder tasks. Multiple students also said that it would be more difficult to compare each other's performance when they got different tasks. Lastly, students thought they all had to do the same task so that the teacher could judge them better, which could suggest that giving all students the same task might emphasise the importance of performance.

Table 10.1

Potential Influences of Fixed-Ability Practices on Students' Motivational Beliefs as Revealed During Student Interviews

Practice dimension	Fixed mindset	Mastery	P.-approach	P.-avoidance	Attribution	Self-concept	Emotion
Tasks							
Unsophisticated task differentiation		Reduction				Reduction	Negative/fair/unfair
Unidimensional tasks		Reduction	[a]Increase	Increase			
Grouping							
Little group and pair work		Reduction	[b]Increase	Increase			
Homogenous seating/ability grouping		Mixed		Mixed		Mixed	Mixed
Stability (school-level)	Increase	Reduction	[c]Reduction	[c]Reduction	Reduced effort attribution		Unfair
Evaluation							
Public evaluation			[d]Increase	Increase		Mixed	Mixed
Performance over process emphasis		Reduction	[e]Increase	Increase			Negative
Little promotion of self-regulation							
Little choice							Mixed
Little promotion of self-regulation	Increase	Reduction				Mixed	
Differential value	Increase	Reduction				Reduction	Negative
Little promotion of risk-taking							
Little encouragement of help-seeking		Reduction		Increase		Reduction	Negative
Unconstructive handling of mistakes	Increase	Reduction		Increase	Reduced effort attribution	Reduction	Negative

Note. Colour coding: green = beneficial effects, yellow = neutral/mixed effects, red = detrimental effects. Effects on performance-approach goals cannot as easily be categorised because such goals can be adaptive or maladaptive depending on how they are expressed:
[a]Detrimental because students referred to appearance concerns.
[b]Mixed because students referred to the importance of performance in general without reference to either normative or appearance goals.
[c]Mixed because students referred to a reduction in general performance pressure without reference to either normative or appearance goals.
[d]Detrimental because students referred to an increase in appearance goals.
[e]Mixed because students referred to an increase in performance pressure without reference to either normative or appearance goals.

Many students disclosed that getting easier tasks made them feel less intelligent or capable, which indicates a reduction in self-concept. One boy said that getting easier tasks would make students think *"that [they] are not, like, as good mathematalics, mathematations* [sic]*,"* and a girl described that *"it might make them feel like they are more stupid than everyone else."* Similarly, some girls explained that getting easier tasks *"can knock down your confidence,"* and that students *"don't feel like [they] are good enough,"* and it *"lower[s] the people's with easy tasks self-esteem."* Unsophisticated task differentiation also seemed to send the message that teachers have lower expectations for some students. One boy suggested, *"if you are not as good to do the harder test then she might not think you are capable of doing that sort of thing."* Another boy said that it would annoy him if his teacher gave him easy tasks because his teacher knew that he was very capable. However, one girl thought getting easier tasks was not unfair but just a sign that these students *"are not actually gonna be, like, meant to be in set A"* and should, thus, move to a lower ability set.

In contrast to teachers assigning different difficulty levels, students of Ms Barns and Ms Green could often choose their difficulty level themselves. Although students reported that this was because they had different abilities, they did not seem to perceive this as something negative but described that it meant you could challenge yourself and improve the level you worked at. They also thought that everyone could at least learn something that way. If students would not have the choice, one girl said, *"you'll be her little puppet."* There was no sense of competition amongst students regarding who worked on what difficulty level, although one boy mentioned he would speed up on easy tasks to be ahead of others.

10.1.1.2 Unidimensional tasks

Very few students linked aspects such as working on tasks that can be solved in different ways to their motivational beliefs or emotions. Some preferred using their own methods because they thought these were faster, while others stated it was better to follow their teachers' way because they were more experienced. Potential links between receiving challenging extension tasks and students' motivational beliefs were revealed. First, a multidimensional extension task (e.g., a task that students can make more challenging by choosing more difficult numbers or using different approaches) might foster students' mastery goals because students stressed that such a task pushes them, *"makes [them] use [their] brain," and "makes [them] think."* Students also said that extension tasks kept them busy. Moreover, a few students reported that these prevented them from comparing each other because they could work on the extension task rather than letting the class know they had already finished previous tasks. By preventing social comparison, these tasks might inhibit performance goals. Importantly, students reported that there was no competition regarding who got on to the extension tasks.

10.1.2 Grouping and students' motivational beliefs

10.1.2.1 Group and pair work

Only a few instances revealed possible links between engaging in group or pair work and students' motivational beliefs. Group and pair work seemed to be related to mastery and performance emphasis in the classroom and might, thus, eventually influence students' goal orientations. Multiple students indicated that pair work helped them to *"give each other ideas"* and support each other, which was particularly true for heterogeneous pairs as discussed in the next subsection. This shows that working with partners might place emphasis on gaining a deeper understanding and, thus, mastery. Moreover, one girl described that they could not engage in pair work because the teacher *"wants us to concentrate and do this so we can get ready for our post-assessments (. …) Like, practice to do this on our own."* This suggests that the absence of partner work might emphasise the importance of individual knowledge and test performance, which might foster performance goals.

10.1.2.2 Seating and ability grouping

Discussions about ability grouping revealed that school-wide ability sets may influence students' self-concept and their perceptions of other students' ability directly, and also indirectly through teachers' expectations. Multiple students held the assumption that students belonging to the same set had the same ability and, accordingly, that students in higher sets were more able than those in lower sets. For instance, one girl argued that they were all in the same set based on their SAT scores, which meant they all had *"the same capability and everything,"* although *"there are some people that might be a tiny bit better than everyone else."* Students also had different expectations for different ability sets. For example, one boy thought that getting demerits for declining performance would happen in lower ability sets even though it never happened in his higher set. And one girl in the highest ability set worried, *"if we don't stay in the same set, then we can't, like, go to our full potential,"* showing that she perceived a lower learning potential for lower sets. Students also noted that teachers had different expectations for different sets. One boy explained that because their teacher was teaching *"the highest set of the year, she, like, she kind of has higher* *expectations"* and *"[n]ow that we have a proper set they kind of, um, they have, like, higher expectations of people or lower."*

Within-class pairings based on student performance caused quite passionate discussions amongst students because some thought homogenous pairs were better, while others argued teachers should form heterogeneous pairs. The former argued that pairing students who have similar SAT scores and who *"have similar thinking and like understand the same amount"* was better because otherwise, it would be more difficult to talk and understand each other. One boy explained:

> *(...) say me and [student] have the same score, if I'm with him, then I think we'd be able to talk about how we are feeling about it. Because we both would be pretty much on the same line of how we've done it.*

This also indicates that students may be more likely to admit that they struggle if their partner works at a similar level, and it shows that students seem to assume that test performance is indicative of ability. In addition, one girl mentioned that it was better to place similar students together because otherwise *"if you are, like, with a confident person, like, besides you, you feel less confident,"* indicating a reduced self-concept and, potentially, a performance-avoidance goal. One girl also argued that homogenous pairs would help because when the teacher explained something to one of the students, the other could always hear it too. Students did not discuss in-depth what they thought about weaker students being sat at the front, but some boys argued that this was a good practice because these students needed more help, which they would not receive at the back.

Students arguing for heterogeneous pairs focused on the benefits for learning. They argued that students who perform differently *"can help each other."* One boy even said, *"I could be sitting with the highest because then he can improve me,"* revealing his belief in growth. Students also thought heterogeneous pairs would mean they needed less help from the teacher, which was good because *"sometimes your friends understand more about what's happening than the actual teachers do."* Overall, it seemed that forming heterogeneous pairs and allowing partner work emphasised learning and mastery and might as such foster mastery goals over time, whereas homogenous pairs were linked to performance concerns.

10.1.2.3 Stability

Discussions suggested grouping stability might influence students' mindsets, effort attributions, and achievement goals. Students associated being able to regularly change school-wide ability groups with being able to grow and improve their weaknesses through effort, which indicates a link to growth mindset and effort attribution. For instance, one girl described that working hard meant you could move to higher ability groups, and another said it would even be OK to move downwards because *"they, like, help you understand what your weakness is,"* which is also consistent with a mastery goal. Similarly, students argued that having the chance to move sets, even downwards, made *"you feel like you're gonna get motivated more and you are gonna move up to the top set."* Nevertheless, students described that flexible ability groups created performance pressure from teachers and parents. One girl admitted:

> *The set changes is the one things* [sic] *that worry me because they are, like, so I've been in set A for most of the time, the whole year and that is really worrying. The thought of going down to B.*

On the other hand, students thought not being able to move was unfair because if *"you're in the lowest set for the whole year and you know, you're, like, you're supposed to be in the top set, it's like you can't, like, you can't necessarily improve."* Overall then, having flexible groups motivated students to move to higher sets and work on their weaknesses but this also caused performance pressure, while having stable sets was considered as unfair because students could not improve.

10.1.3 Evaluation and students' motivational beliefs

10.1.3.1 Public evaluation

Discussions with students about teachers' use of public evaluation exposed potential effects on students' performance goals and self-concept. There were many consistent examples in which students described that they were happy to have the class know how well they were doing, that they started bragging, and wanted to be recognised for good performance. Many students reported that if the teachers asked students to indicate who got something correct, it *"gives you, the people that got it right, like a chance, like, brag about it, like, not actually like directly brag about it but, like, you, like, put your hand up really high."* Others also said, *"people start being boastful."* Students in the highest ability set actually thought that *"there has to be competition. You have to compete with others"* and that *"rivalry is… like a boost."* Getting merits in front of the class was considered to be fine by most students, and it even motivated them to work harder if they saw that others got praised. However, students who received no praise did not *"feel appreciated for what they've done,"* and students thought it *"kind of feels a bit, like, sad that they didn't really get any gratitude for it."* This indicates that public praise can foster performance-approach goals.

On the other hand, students *"feel like really humiliated"* and embarrassed when the class knew they were struggling. Students explained, *"sometimes people don't want to be, like, have the class knowing that they are confused about things 'cause other people might find it easy."* Over time, public evaluation, such as sharing students' test results or asking them how many correct answers they got, can thus foster performance-avoidance goals. Public evaluation could even lead to negative reactions from other students, which can discourage students from participating in the class. One boy described:

> *(…) there are people that can get really upset about things that I've seen and, um, if they get, like, um, kind of bullied around about it, I think it can be bad, like, it can ruin their morale about, well actually being in the maths class and having a go. It means that, um, that they will be more scared to have a go at answering questions that Missus asks.*

Many described that public evaluation could also lead to cheating behaviour as one girl admitted, *"once I put [my hands] up when I didn't actually get all right because I didn't really want to be left out."*

Besides relationships with performance goals, many students revealed that public evaluation may affect their self-concept either positively or negatively depending on how they were doing. Public evaluation gave lower-performing students *"feelings that they are not really good at maths,"* especially when they saw that others performed well. In contrast, those that were doing well felt *"proud."* It also seemed to affect how students thought about each other as one girl declared:

> *I know who's the smartest of the class (. …) it's half of it is from when she asks who got them all right (. …) you don't even have to turn around to see. You know they put their hands up.*

One boy even exclaimed, *"they are too dumb, so they are not getting it right."* Some students also revealed that seeing others getting praised made them question their own maths ability. One girl explained, *"she says it to, like, some people more than others. So, they might think, 'oh am I not doing good? Um, am I then rubbish at maths?'"* The students thought it was unfair that teachers sometimes pointed out how students were doing as one boy described:

> *I kind of found it kind of unfair of saying that 'cause you are like saying, 'oh look at these guys' (. …) I'd rather her not to tell anybody what everybody else is doing so you don't have to worry.*

Students also found social comparisons annoying *"'cause you feel like, um, pressured that you need to do more"* and said you might start to rush and then make unnecessary mistakes. Some also suggested the weaker students might feel sad, but not all agreed.

Notably, some students reported that they did not mind public evaluation and that they did not perceive any competition because of it. One girl explained, *"I don't mind because, like, not everyone is at the same pace with, like, learning maths."* And another said *"well I'm not like a, like a jealous person. So I would say that I'm proud of that person because that person worked hard."*

10.1.3.2 Performance over process emphasis

Beginning with performance emphasis, students at Meadows School, where formal assessment was very prevalent, reported that the emphasis on test performance created much performance pressure and stress. Additionally, performance emphasis indicated to students that they needed to learn in order to perform well, rather than to gain mastery. Some students thought that teachers cared mostly about their test performance rather than real understanding because teachers often pointed out that they had to revise for tests. One said their teacher thought *"it doesn't matter how hard we try, like. It, um, like, it's more, like, what you get on your post-test."* Some students even assumed that they were not allowed to work in pairs because they needed to do the work on their own during tests. Moreover, some students accepted that the teacher had not enough time to help all students understand something *"because she wants to get through this term really quick so that at the end, we have enough time to revise."* Students in these classes also thought *"it's very crucial that you go [to maths clinic]. So that you can, like, improve on your scores."* All these instances suggest that these students came to believe that they needed to learn to perform well rather than to master the material, which indicates the adoption of performance goals and a reduction in mastery goals. Performance emphasis might also reduce students' self-esteem, as one girl explained:

> *I wouldn't really feel I'm worth her time if she, if she is only caring about what I do, if it's wrong or right. But if she cared whether I was trying hard or not I feel like I'm worth the time, like, she's put, she's put into, er, helping me learn.*

Notably, some students at Meadows School did believe that their teachers cared more about effort than performance. And some students pointed out that because of the regular tests, *"we can see how much we've improved,"* which might show some students that they can grow.

Speed emphasis was also mostly negatively perceived by students although some said it *"pushes our brain"* and *"alerts us more and, like, makes us think quicker,"* which would indicate learning gains. One girl even argued that speed emphasis was good because of her performance concerns. She explained, *"if there's this one person who hasn't finished, then it's holding the class behind. And so you can't, like, get on with the lesson,"* to which another replied, *"you might be behind other schools."* In contrast, many thought that speed emphasis could be humiliating for some students or just annoying because *"someone hasn't got enough time to finish"* and might make unnecessary mistakes *"while [they] could have gotten it right in [their] own time."*

Students interpreted teachers' product-oriented feedback, such as criticism for not solving enough tasks, differently, seemingly depending on their attributional style. Some considered being criticised for poor performance as fair *"'cause that's obviously showing that you, that you didn't, um, gave, revise,"* while others thought criticism for poor performance was unfair because *"everyone learns at a different pace and we learn different things in a certain amount of time. So you can't really blame us for not learning it."* Moreover, some students said it was unfair because students had too much to do, so that they did not have enough time to revise properly. Accordingly, students perceived product-oriented criticism differently depending on whether they made effort, ability or situation attributions, which was also found for product-oriented praise. Getting praised for good work was not consistently linked to students' motivational beliefs, but one boy mentioned that it could lead to cheating behaviour, which indicates that it may foster performance goals. However, there were too few instances to support this.

Turning to process emphasis, students consistently linked having to explain how they got their answers and show their workings to the importance of understanding and learning and, thus, mastery. If teachers asked students to explain their answers, students thought they wanted to ensure that students understood what they were doing. In addition, it helped other students in the class as one student described, *"we can see how the person worked it out and if their method is better for us."* Moreover, many instances showed that students linked process-oriented praise, especially praise for improvements, to the belief that they could indeed improve and *"that shows that [the teacher] wants us to try."* This motivated them to *"work even harder."* Process-oriented praise also showed students who were not getting praised that they needed to work harder. Lastly, all students valued being challenged as they thought they would otherwise not be learning. Students thought that challenges were *"what makes [them] really think."* If teachers would not do this, the lessons would be *"really easy and boring"* and they *"wouldn't be learning anything."* Hence, if students would not feel challenged, they might come to believe that learning was not as important to their teachers, which could reduce mastery goals.

10.1.4 Autonomy and students' motivational beliefs

10.1.4.1 Little choice

Interviews did not reveal any notable relationships between students having little choice and their motivational beliefs. Students' emotional responses to little choice varied. Some students did not like having little choice because they thought that doing things their own way would be easier and more efficient. Others thought it was justified that teachers made all decisions because they were more experienced than the students. One boy also said his teacher gave them little choice because *"[s]he wants us to succeed (. ...) Wants us to do more to progress our learning,"* indicating that he thought students tended to make "lazy" choices, while the teacher pushed them more.

10.1.4.2 Little promotion of self-regulation

Interviews showed that the promotion of self-regulation might foster students' growth mindset and mastery goals. Some students discussed that their teacher wanted them to assess their own or each other's work *"because she wants you to look at your own mistakes"* and *"if you check it then you have, like, more understanding."* This indicates students thought self-assessment was important for learning. This learning emphasis might over time foster their mastery goals. Students further discussed that self-assessment showed that their teacher *"wants us to be teachers"* and that *"she won't have to help you as much so you can be more independent on it,"* which reveals a link to students' feelings of autonomy, although this was not one of the constructs of interest in this study. Notably, students at Meadows School also pointed out that many of them cheated when they self-assessed because their scores sometimes determined how many merits they got.

Only the students at Woodpark School were asked to evaluate their own learning besides engaging in self-assessment. There were only a few instances that revealed potential links between self-evaluation and students' motivational beliefs, and these were to mastery goals and a growth mindset. One boy explained they had to self-evaluate *"so that we, we could improve on what, on that next lesson and then, like, if we know what went well, we could, like, look back from that and say, 'well I can be better now.'"* Similarly, a girl described that they had to identify *"something you need to improve."* Other than this, students described self-evaluations as helping teachers to understand what students struggled with and, thus, to decide what they had to teach more.

Students at Meadows School discussed their teachers giving them targets for their test performance, and they had mixed opinions about this practice. One boy was concerned that the teacher might set targets that were too high and that they should get *"more of a chance."* In contrast, one girl thought this was good because students would give themselves targets that were too low. Importantly, one girl said, *"I feel like she has got a pretty good idea of, like, what she, what target we should meet."* This indicates that she trusted the teachers' judgement, which seems fine if the teacher had high expectations but could become problematic if they set relatively low goals. If students think their teachers set the goals appropriately, receiving a low target or goal might reduce students' self-concepts. However, there were too few instances to draw strong conclusions about this.

10.1.5 Differential value and students' motivational beliefs

Discussions about differential value, mostly regarding whether or not teachers encouraged all students to participate, revealed that students perceived an emphasis on learning and the potential to grow if teachers engaged all students. In contrast, differential value seemed to lower students' self-concept. Students thought that teachers asked all students to participate in order to engage them and make sure that all focused on the lesson to help them learn. For instance, one boy described, *"she picks pupils where she thinks, 'oh they might not know it yet, so I can help them,"* and students thought the teacher *"will show them where they've gone wrong."* This emphasised the importance of, and the potential for, learning because teachers did not give up on "weaker" students. Similarly, one boy explained how important it was for students' learning that they were asked to give answers even if they struggled because *"if you answer questions, then you remember them more… it'll stick in your head for longer."* If the teacher called on students who found the work easy *"it won't stick as well and you forget it after a while."* Over time, this emphasis on learning and the potential for growth might foster students' mastery goals and growth mindset.

All students reported that their teachers tended to frequently call on students who were a bit weaker, shy, or did not participate as much. Depending on who the teachers called on, students learned about what expectations their teachers had for them. For instance, one boy described, *"they are really smart, and Miss doesn't choose them because they know the answer (. …) So she kind of goes to people who don't know a lot. For example, um, she goes to [student] quite often, not a lot."* One might expect that this could lower the self-concept of students who were perceived as weaker, but students instead thought that giving an answer in front of the class would *"boost their confidence."*

Teachers sometimes called on the stronger students to move on the lesson more quickly. Students viewed this very negatively although one student mentioned that the stronger students could teach others this way. One boy indicated that calling mostly on stronger students could negatively affect weaker students' self-concept as he described, *"I don't like it at all. Because then it looks like she doesn't believe in us. And she doesn't think we can give the answ-, answer."* Another explained he would be unhappy and would feel *"underpowered in a way. Like, she ... you are trying to have a go at it. She just goes straight away to the person who knows it."*

10.1.6 Little promotion of risk-taking and students' motivational beliefs

10.1.6.1 Little encouragement of help-seeking

Discussions showed that whether or not teachers encourage help-seeking might affect students' goal orientations and, potentially, their self-concept. Students thought that their teachers cared for them and wanted to help them learn if they could ask their teachers for help and if teachers answered their questions thoroughly. Hence, encouragement of help-seeking appears to emphasise mastery, which might foster students' mastery goals over time. Students pointed out, *"they are teachers, they are supposed to help us learn and if we can't, if we can't ask them questions, then how are we supposed to learn?"* Although not said explicitly, this indicates that teachers who do not encourage help-seeking might convey to students that they do not care as much whether or not students learn, which might reduce students' mastery goals over time. Moreover, encouragement of help-seeking may inhibit performance-avoidance goals as one boy described, *"I don't mind telling her about the question. And I just don't really mind her knowing about that I didn't get the work."* In contrast, not being allowed to ask for help seemed to emphasise performance as one girl thought that this was because *"sometimes you will just have to, er, solve your ow-, your problems yourself. 'Cause you need to learn these things before the post-assessments starts."* Discussions revealed that students' own performance goals may influence whether they ask for help but that, nevertheless, teachers' reactions to help-seeking could change their performance goals. When one girl admitted, *"I don't really like asking (. ...) Then I know she knows that I'm struggling,"* another responded, *"I used to feel the same way (. ...) I was scared to ask my teacher (. ...) But now I'm, like, actually OK to ask her because she's made me feel OK to ask her things."*

One boy even said that not being able to ask his teacher for help would make him feel *"kind of dumb,"* suggesting negative effects on students' self-concept, but there was not enough evidence to draw strong conclusions. Generally, students showed very negative emotions to teachers not encouraging them to ask for help, such as being *"upset"* and *"angry,"* and feeling as though the teacher did not take them seriously. Students' perceptions of teachers' encouragement of peer support have already been addressed in the subsection on group and pair work and are thus not repeated here.

10.1.6.2 Unconstructive handling of mistakes

Student interviews showed that teachers' unconstructive handling of mistakes, such as quickly moving on to other students, were perceived very negatively and considered useless. It also seemed to negatively affect students' self-concept. In contrast, adaptive reactions to mistakes, such as using mistakes as learning opportunities, seemed to emphasise the importance of learning and the possibility to grow through effort, which is likely to foster mastery goals, a growth mindset, and effort attributions over time.

The most common student reaction to teachers swiftly moving on after mistakes was *"that's useless"* because *"if you don't know how you've done it wrong, you don't know how to learn from it."* Two girls discussed that their teacher sometimes wrote down the correct answer and that they then had to copy it *"so we understand in the future."* However, they said some students are then thinking *"'what am I doing just writing it down' (. ...) because they might not understand how they got it."* Feeling as though they *"haven't learned anything"* and that their teacher does not care whether everyone understood might adversely affect students' mastery goals over time. Additionally, moving on after mistakes seemed to negatively affect students' self-concept as some reported, *"[it] makes me feel not as intelligent"* and *"it makes you feel dumb."* They also said it was *"gonna make the child feel bad"* or *"a bit down"* and as though the teacher was disappointed, although not all students agreed. Notably, some students at Meadows School, where testing was very prevalent, argued teachers sometimes had to move on even if it made students feel bad because they would otherwise lose time that they needed for test revisions, whereas others argued *"this information might be really important to the test."*

When teachers gave students the feeling that it was OK to make mistakes because everybody made them and that these were important learning opportunities, students started to see the benefits of mistakes. For instance, they regarded making a mistake as something that was *"making you improve. 'Cause if you make a mistake then you can learn from that mistake."*

Students also described, *"if you make a mistake it helps* [the teacher] *to know where you are going wrong and where everyone else is going wrong."* Similarly, one girl described a poster that she saw in the school, and now also had in her own home, that portrayed mistakes as an important part of the learning process. She explained:

> *(...) that inspires me because it shows me, like, how I sh-, so if I get a mistake, then it doesn't really matter. But then if you try again and make mistakes, then you have to keep on trying until you get it right.*

This again highlights that students focused on the growth potential and that they believed effort could lead to such improvements. Some students also said that they knew their teacher valued effort more than performance because she did not mind mistakes but asked them to try harder next time.

Similarly, discussing mistakes demonstrated that students could learn from mistakes as one girl described that the teacher often held up mistakes and then explained what went wrong, which *"might be helpful for others because they might not under-, they might have wrote* [sic] *the same thing as well. So when they go over it, then they might understand it a bit more."* Moreover, students stopped worrying about others knowing they made a mistake as a student explained, *"they don't feel bad that it's only them making that mistake (. ...) It helps other people too."* This indicates a reduction in performance-avoidance goals but there were too few instances to draw strong conclusions.

Students highly valued the opportunity to self-correct mistakes and to try to improve, which showed them that their teacher *"wants [them] to improve on [their] intelligent* [sic] *in maths."* Opportunities for self-correction also seemed to foster mastery goals while inhibiting performance-avoidance goals as one student described:

> *(...) if we get it wrong then she says, 'have a look over that again.' We just like try and do it again. Like we don't really think, 'oh no, I've got it wrong and they've got it right. I'm so sad and everything.' Just you, you are more focused on doing it for yourself and everything.*

10.1.7 Summary on students' perceptions of FAPs

As expected, discussions with students revealed that many of the FAPs seemed to be negatively related to their motivational beliefs, although I do not want to make causal claims because of the small number of interviewees. Teachers' unsophisticated task differentiation might lower students' self-concept because students reported they would feel less competent if they got easier or less work than others. It might also lower their mastery goals over time because students thought easier tasks provided fewer learning opportunities, which might convey that learning is not as important. In contrast, students perceived being able to choose their difficulty level positively. Discussions did not indicate that unidimensional tasks might adversely affect students' motivational beliefs other than that the provision of challenging extension tasks, which is an aspect of multidimensional tasks, might foster mastery goals and reduce performance goals. Students thought extension tasks pushed them and prevented them from seeing who finished tasks early, thereby reducing social comparison. These relationships were only tentative though as they were not supported by many students.

Students related engaging in group or pair work to a learning emphasis, while they thought not engaging in pair work was due to performance concerns. Accordingly, these practices could potentially influence students' goal orientations over time. Perceptions of within-class and school-wide ability grouping were very mixed. School-wide ability sets seemed to influence students' self-concept and their perceptions of other students' ability either positively or negatively depending on their set. Some students argued that homogenous pairs would benefit learning because students with similar test performance thought in similar ways and could receive teachers' help simultaneously, while others thought heterogenous pairs benefitted their learning because students could learn from each other. Hence, both pairing strategies could potentially foster mastery goals depending on students' perceptions. Students preferring homogenous pairs appeared to have greater performance-avoidance goals, but the direction of influence is unclear. Flexible ability groups seemed to show students that they can change their ability through effort although it also created performance pressure. Students thought that stable ability groups gave them fewer opportunities to learn and improve.

Public evaluation appeared to foster students' performance goals by either embarrassing students so that they would shy away from future participation or by encouraging them to brag about their performance. It also seemed to either increase or reduce students' self-concept depending on how they performed. Test and performance emphasis appeared to give students the feeling that they learned to perform well, rather than to gain mastery. This indicates an adoption of performance goals and a reduction in mastery goals. How students interpreted product-oriented feedback seemed to depend on their attributional style, while process-oriented feedback was consistently linked to the importance of learning, which might foster mastery goals over time.

Students revealed no notable relationships between having choice and their motivational beliefs. Yet, promotion of self-regulation might foster growth mindsets and mastery goals because students thought that having to assess or evaluate their work would help them to learn by showing them how to improve. There was also some indication that the targets that teachers set for students might be related to students' self-concept because they seemed to trust teachers' judgements.

Not showing differential value but encouraging all students to participate seemed to emphasise the importance of learning and the possibility of improvement, which might foster mastery goals and a growth mindset. In contrast, differential value seemed to reduce students' self-concept because students appeared to infer teachers' expectations based on how they were being treated compared to others.

Encouragement of help-seeking seemed to emphasise the importance of mastery and give students the feeling that it was OK to show that they struggled, which suggests a reduction in performance-avoidance goals. Not encouraging help-seeking might reduce students' self-concept as one student explained that not feeling able to ask for help made him feel dumb. Teachers' constructive handling of mistakes appeared to emphasise the importance of learning and the possibility to grow with effort, which is likely to foster mastery goals, growth mindsets, and effort attributions. Knowing that it was OK to make mistakes also seemed to reduce students' performance-avoidance concerns. In contrast, unconstructive handling of mistakes, such as moving on to another student, was perceived as useless because students could not learn anything, which might reduce mastery goals. This also gave students the feeling of not being as intelligent, indicating a reduction in their self-concept.

Chapter 11

Discussion

Decades of research have demonstrated the importance of students' mindsets for their learning (Burnette et al., 2013; Dweck & Leggett, 1988; Dweck & Molden, 2017). Mindsets can give rise to a meaning system consisting of interrelated motivational beliefs including goal orientation (Payne et al., 2007), achievement attribution (Blackwell et al., 2007), and ability self-concept (Gunderson et al., 2017). In addition, they even impact achievement (e.g., Blackwell et al., 2007). Understanding how mindsets develop is thus a great concern for educational psychology (see Haimovitz & Dweck, 2017). This study's purpose was to understand the contribution of fixed-ability practices (FAPs) – teaching practices that are based on the notion that ability is fixed – to the development of fixed mindsets and related maladaptive motivational beliefs. The first aim was to determine what FAPs are. I developed a conceptualisation and a corresponding measurement tool, based on student and teacher reports, and then examined the dimensionality of student-perceived FAPs (RQ 1). The second aim was to understand why teachers engage in FAPs. Thus, I examined the relationships between teachers' own motivational beliefs and their engagement in student-perceived FAPs (RQ 2), and I explored teachers' justifications for engaging in these practices through interviews (RQ 4). Note that FAPs were only characterised based on the students' perspective but not teachers' perspective as the teacher scales had low internal reliability. Lastly, the third aim was to understand how FAPs affect students' motivational beliefs and achievement. Hence, I examined whether student-perceived FAPs predict students' motivational beliefs and achievement (RQ 3), and I explored how students perceive these practices, in particular in relation to their beliefs (RQ 5).

Results showed that student-perceived FAPs comprise five dimensions at the individual level: unsophisticated task differentiation, public evaluation, little promotion of self-regulation, differential value, and little promotion of risk-taking. There was no overarching FAPs dimension that represented all these practices well. At the classroom level, there was one uniform dimension. Corresponding FAPs scales for teachers had inadequate psychometric properties and were thus not used in subsequent analyses. Correlations between teachers' motivational beliefs and students' shared perceptions of FAPs were all non-significant. Hence, these statistical relationships were not examined further. Thematic analysis of teacher interviews showed that teachers provided multiple justifications for their practices involving beliefs, goals for students, practical considerations, and school policies. Engagement in more adaptive practices was usually explained by adaptive beliefs and values, while maladaptive practices were explained by maladaptive beliefs and values. Yet, teachers rarely referred to their mindset or goal orientation, and relationships between beliefs and practices were at times inconsistent. Longitudinal analyses indicated that students' individual perceptions of teachers' differential value positively predicted fixed mindsets and negatively predicted self-concept. Moreover, perceived little promotion of risk-taking positively predicted performance-approach goals. In addition, there were several notable trends. At the classroom and teacher level, students' shared perceptions of FAPs positively predicted performance-approach goals. Discussions with students indicated that they perceived FAPs negatively and that FAPs may foster maladaptive motivational beliefs. Moreover, students' initial beliefs seemed to affect how they perceived some practices. The qualitative and quantitative results were mostly aligned, but student interviews revealed additional potential influences of FAPs on motivational beliefs.

This chapter provides a discussion of the main results of the study. The three aims are addressed in turn by synthesising and evaluating the quantitative and qualitative findings (Sections 11.1 to 11.3). The following chapter addresses the contributions and implications of this study and considers its limitations as well as the arising avenues for future research.

11.1 What are fixed-ability practices?

The first aim of the study was to develop a conceptualisation of, and corresponding questionnaire for, FAPs, and to assess the dimensionality of student-perceived FAPs. Quantitative results based on student reports showed that FAPs comprise five interrelated but distinct dimensions on the student level, namely unsophisticated task differentiation, public evaluation, little promotion of self-regulation, differential value, and little promotion of risk-taking (see Section 5.2.1). On the classroom level, FAPs formed one uniform dimension (see Section, 5.3). The framework and final questionnaire thus include a broad range of practices that are rooted in theory and previous research. As such, it is a comprehensive measurement instrument that provides a holistic view of student-perceived FAPs in the classroom. Developing a questionnaire was particularly important given that previous mindset teaching frameworks have been examined using only qualitative data. Hence, no comprehensive measurement tool was previously available that enabled researchers to simultaneously examine this range of distinct but interrelated practices.

The factor structure, and thus final scales, were based on student reports alone because the teacher sample size was too small for conducting a factor analysis. Unfortunately, the corresponding teacher scales did not reach adequate internal reliability (see Section 5.5). Although students and teachers responding to items in different ways cannot be excluded, the more likely explanation is a too small teacher sample size. Common rules of thumb suggest at least a sample size of between 200 (Kline, 1986) and 400 (Charter, 2003) because small samples can lead to imprecise estimates of the population reliability.

The final dimensions of student-perceived FAPs were mostly consistent with the initial conceptualisation grounded in the literature, but there were also important changes. Four dimensions that I conceptualised (unidimensional tasks, ability grouping, grouping stability, and little choice) were not retained in the final questionnaire. From a theoretical perspective, multidimensional tasks are in line with teaching for a growth mindset because they allow students to show their skills in different ways, whereas unidimensional tasks may foster fixed mindsets (Boaler, 2013; Sun, 2018). However, items assessing this construct were not sufficiently correlated (see Section 5.1), indicating that this might be too broad a concept to be captured in a single dimension. Items referred to diverse practices such as setting tasks that can be solved in different ways, providing extension tasks, and providing a range of different learning activities. The assessment of this dimension thus needs to be considered further in future research. Ability grouping and stability were also not part of the final scale. Grouping items were excluded because they assumed students would engage in group or pair work, which was often not the case. Surprisingly, many students responded regardless, which indicates the need to formulate items in a way that ensures that only students who engage in pair or group work complete them (see Section 4.1.3). These items should be revised because within-class ability grouping is theoretically linked to notions of a fixed ability (Marks, 2013). Interestingly, it has been found that British secondary school teachers often allow students to choose their group or seating partners, but that students predominantly decide to work with friends of a similar achievement level, which thus leads to homogenous groups (Kutnick et al., 2005). This indicates that teachers might have to actively encourage heterogeneous seating and group compositions. The little choice scale was also not retained because the items, adapted from Assor et al. (2002), were not sufficiently correlated to be considered for factor analysis (see Section 5.1). Although items were only slightly adapted to simplify wording, some items were excluded during piloting, which might have affected the scale's reliability. Hence, the scale needs to be developed further.

Some dimensions that I initially conceptualised as related but distinct merged into a single dimension (see Section 5.1). First, performance over process recognition and social comparison merged into the dimension of public evaluation. This suggests that teachers who publicly evaluated students' performance also tended to publicly recognise students' effort and progress, at least as perceived by students. Other possible explanations include that students did not distinguish effort and performance evaluations or that the items were ambiguous. For example, the item "My teacher talks about how well students do in maths in front of the whole class" could refer to both effort and performance. The key aspect was that teachers made evaluations in front of the class and, thereby, provided cues for social comparison. Notably, three of the five public evaluation items were excluded from the aggregated, uniform FAPs factor because of low loadings. A likely explanation is that teachers who engage in public evaluation do not necessarily also engage in the other fixed-ability practices. This was supported by low correlations between student-perceived public evaluation and little promotion of risk-taking as well as self-regulation. Similarly, previous research found that teachers do not seem to consistently engage in either adaptive or maladaptive practices related to the communication of mindset (Sun, 2019). Second, the dimension little promotion of risk-taking was initially hypothesised to be comprised of the two subdimensions little encouragement of help-seeking and unconstructive handling of mistakes but these subdimensions merged into one. This demonstrates their close link. Both practices relate to whether it is OK for students to take risks and reveal that they struggle (Stipek et al., 1998) or whether making mistakes and asking for help is considered a sign of low ability and, thus, something that needs to be avoided (Dweck, 1999; Karabenick, 1998).

The study also examined whether there was an overarching FAPs dimension at the student level. Although the bifactor model, including such a general factor, did not fit worse than a simpler model, the general factor did not represent all items well (see Section 5.2). In particular, several items from the unsophisticated task differentiation and public evaluation scales did not load strongly on the general factor (see Table G.3). This is similar to the uniform FAPs dimension at the classroom level (see Table 5.4), which is why some items were excluded from that dimension when examining the effects of students' shared perceptions of FAPs. Hence, the simpler model without an overarching FAPs dimension was used to represent students' individual perceptions. These findings suggest that teachers who engage in a given FAPs dimension do not necessarily also engage in another. Similarly, Sun (2019) found that teachers do not consistently engage in practices that are related to either a growth or a fixed mindset. For instance, one teacher in her case study used comparative structures, which is associated with a fixed mindset, while providing students with extra opportunities for help, which is more in line with a growth mindset. This indicates that attention must be given to each specific practice dimension in terms of why teachers engage in that practice and how this could be avoided.

The FAPs dimensions support and extend existing frameworks of teaching practices for mindset. Previous frameworks also included practices related to public evaluation and social comparison (Rissanen et al., 2019) as well as promotion of risk-taking, such as handling of mistakes or encouragement of help-seeking (Boaler, 2013; Rissanen et al., 2019; Sun, 2018). Promotion of self-regulation is not included in these teaching for mindset frameworks despite its theoretical and empirical links to mindset; Ommundsen (2001) found that provision of choice predicted growth mindsets, presumably because it elicits feelings of control over the learning process. Moreover, teachers with fixed mindsets tend to provide less autonomy support (Leroy et al., 2007). Not included in any of these frameworks is teachers' differential value. This is surprising because a prior qualitative study showed that providing encouragement and support for high achievers appears to be more characteristic of teachers with a fixed mindset (Rissanen et al., 2018). As discussed below, this dimension was particularly detrimental to students' motivation, which highlights the need to include it in the conceptualisation and future research.

The current conceptualisation challenges Rissanen et al. (2019) who argued that task differentiation is part of a growth mindset pedagogy. It is unclear from their paper how task differentiation was defined, that is, whether or not it went beyond setting fewer or easier tasks for lower-performing students. Interestingly, Rissanen et al. (2018) proposed that teaching to ability, such as lowering task difficulty for lower-performing students, is part of a fixed-mindset meaning system, which seems to contradict their later claim that task differentiation is based on a growth mindset (Rissanen et al., 2019). Others have also argued that task differentiation, such as teaching more difficult topics to only some students, communicates fixed mindset messages (Boaler, 2013). What seems to be crucial is how task differentiation is operationalised. The qualitative analysis of interviews suggests that it might be an important difference whether students can choose which tasks to work on or whether the teacher prescribes the difficulty level. The former seems to convey that students can improve, whereas the latter appears to convey differential teacher expectations.

The FAPs conceptualisation and questionnaire also share important similarities to other teaching practice conceptualisations that were not developed from a mindset perspective.

First, they are similar to the ability-based practices framework (Bohlmann & Weinstein, 2013) as already reviewed (see Section 2.2). This is not surprising given that both conceptualisations focus on practices that emphasise student ability. Yet, no quantitative measurement tool existed for their framework. FAPs are also similar to the TARGET framework (e.g., Epstein, 1988) and its corresponding measurement tool (Lüftenegger et al., 2017), but the two can also be distinguished in important ways, which highlights that a new measurement was warranted. On first glance, the frameworks appear almost identical because the naming of the dimensions is very similar[19]. This, however, is misleading because the dimensions are operationalised quite differently. For instance, the task dimension by Lüftenegger et al. (2017) appears to be a combination of multidimensional tasks and promotion of self-regulation, but it does not assess task differentiation. Moreover, their grouping dimension assesses whether students work in groups but not whether these groups are based on ability, and only one out of six of their evaluation items touches on public evaluation. Importantly, the TARGET framework does not include differential value, which was a particularly influential dimension in the current study. A few items related to promotion of risk-taking are included, but these are distributed across the dimensions evaluation and time. Given that this dimension appeared to be very important in the current study, assessing a distinct risk-taking dimension seems to be warranted when interested in influences on students' mindsets. It is perhaps not surprising that similarities between the frameworks exist because of the close relationship between students' mindsets and goal orientations (for a review, see Payne et al., 2007; Schwinger et al., 2016). Nonetheless, these motivational beliefs are theoretically distinct, and the apparent differences between FAPs and TARGET practices clearly show that taking a mindset-lens can direct our attention to important practices that were not considered on the basis of motivational goal theory alone.

11.1.1 Brief summary on what fixed-ability practices are

This study showed that student-perceived FAPs comprise five dimensions, namely unsophisticated task differentiation, public evaluation, little promotion of self-regulation, differential value, and little promotion of risk-taking, at the student level and one uniform dimension at the classroom level. An overarching FAPs dimension at the student level did not represent all practice items well, which indicates that teachers who engage in some FAPs do not necessarily also engage in others. The new conceptualisation substantiated, extended, and challenged existing frameworks. Importantly, the study also developed a questionnaire for the

[19] TARGET stands for task, authority, recognition, grouping, evaluation, and time.

assessment of FAPs. This provides a promising starting point for gaining a fuller picture of these practices, at least as perceived by students, in the classroom, although some scales need further refinement. The contributions to the literature that these findings make are further discussed in the Conclusion Chapter (see Section 12.1.1).

11.2 Why do teachers engage in fixed-ability practices?

After having established the dimensionality of FAPs, at least as perceived by students, this study sought to understand why teachers engage in such practices. Of particular interest was whether their own motivational beliefs could explain their use of FAPs as well as how they justified these practices. To examine the former, I looked at correlations between teachers' motivational beliefs and students' shared perceptions of the uniform FAPs dimension. The uniform rather than individual dimensions was used because these relationships exist at the teacher level. The quantitative results showed that there were no significant correlations (see Section 7.1.1). This can likely be explained by the small sample size of 25 teachers. Consequently, I cannot draw conclusions from this quantitative data but will only discuss relationships between teachers' beliefs and their engagement in FAPs based on the justifications that teachers provided during interviews. In places, I also use student interviews as a means to corroborate teachers' claims about the practices they engage in.

Teachers engaged relatively infrequently in most of the FAPs according to their own reports and student interviews (see Section 9.2.1). Generally, the more-fixed mindset teachers seemed to engage more frequently in maladaptive practices, such as allowing little pair work, using public evaluation, and providing little choice. Nevertheless, the less-fixed mindset teachers also did not consistently avoid FAPs. I asked teachers "why" they engaged in different practices and, thus, elicited spontaneous justifications rather than directly probing for specific motivational beliefs. Teachers rarely provided a single reason for engaging in a given practice. As discussed further below, they gave multiple explanations involving beliefs, values, practical considerations, and school policies (see Section 9.3). Engagement in more adaptive practices was usually explained by adaptive aims, beliefs or values, such as fostering student mastery, while maladaptive practices were explained by maladaptive beliefs or values, such as concerns for students' performance, as well as practical considerations. This supports the wider literature showing that teachers' beliefs are related to the practices they engage in, even if these are not always straightforward (e.g., Cross, 2009; Pajares, 1992; Thompson, 1992). In a relatively recent review, Fives and Buehl (2012) identified three functions of beliefs that may explain why they affect teaching practices. First, beliefs can act as filters for incoming information so that they, for instance, shape what teachers learn about different teaching practices. Second, beliefs can influence how teachers frame a given situation or task, such as what they consider to be good instruction. Third, beliefs can guide teachers' intentions and actions, such as the goals they set and how they are trying to achieve these.

Yet, teachers rarely explicitly explained their practices with their mindsets. This might be because mindsets are implicitly held beliefs (Dweck, 1999). When searching for explanations for engaging in certain behaviours without being explicitly asked about or probed for such implicit beliefs, these beliefs may not be easily accessible to teachers. Teachers could be unaware of their own beliefs or they may simply not have the language to express what they believe (see Fives & Buehl, 2012). Similarly, rather than referring to their own goal orientations, which are perhaps quite abstract, teachers often explained what they valued, and thus aimed to foster, in their students. Justifying why they engage in a given practice by explaining how they hope the practice will affect students seems more straightforward than focusing on their own underlying achievement goals. While achievement goals are thought to be accessible and brought into consciousness, it is not entirely clear when and how these goals can be brought into awareness (see Pintrich, 2000). The current findings suggest that more immediate goals for student outcomes might have been more accessible. Moreover, in line with previous studies (e.g., Raymond, 1997; Spruce & Bol, 2015; Sun, 2019; Wilkins, 2008), current findings also indicated that beliefs and practices are not always consistent, which has important implications for teacher education and professional development, as discussed in more detail in the next chapter (see Section 12.1.5).

It is important to note again that teachers were grouped as more-fixed or less-fixed mindset teachers based on their questionnaire responses. Ms Barns and Ms Green very strongly rejected a fixed mindset in questionnaires, but they both exposed some fixed-mindset thinking during interviews. Similarly, Ms Evans, who was classified as a more-fixed mindset teacher compared to others in the sample, did not reveal any fixed-mindset beliefs when interviewed. This, of course, does not necessarily mean that she did not hold such beliefs but rather that these were not activated during the interview or even that she deliberately chose not to expose such socially undesirable beliefs. Yet, these findings show the complexity of beliefs as well as potential inadequacies in their assessment. All teachers scored relatively low on fixed mindset, which might be because they responded in socially desirable ways. That is, it might have seemed socially unacceptable to say that students cannot change their maths ability, especially given that they were maths teachers. It appears relatively easy to reject explicit fixed-ability beliefs in questionnaires, whereas these might be more difficult to conceal during interviews because they can be implied in teachers' thinking. In line with this, Sun (2019) found that teachers who reported to have a growth mindset in questionnaires revealed some fixed-mindset beliefs during interviews. For clarity, I retained the initial classification based on questionnaires. This seemed appropriate given that I did not classify teachers as fixed and growth mindset teachers but rather as more-fixed and less-fixed mindset teachers, which is thus a matter of degree rather than kind. I will now discuss the justifications for each practice dimension in more detail. Note that I address all dimensions and their subdimensions, including those that were not retained in questionnaires, such as unidimensional tasks, grouping, and little choice.

11.2.1 Tasks

The expectation that more-fixed mindset teachers would be more likely to assign less or easier work to lower-performing students was not supported because none of the four teachers engaged in unsophisticated task differentiation according to their own and students' interviews. Yet, it should be kept in mind that none of the four teachers strongly endorsed a fixed mindset. Teachers' justifications for not engaging in this practice varied (see Section 9.3.1.1).

Interviews indicated that having high expectations for all students and believing in growth may explain not using unsophisticated task differentiation. For instance, Ms Evans and Ms Barns did not want to mandate which students worked on what tasks because this would set up different expectations. Ms Barns also explained that she did not assign different tasks because a student can sometimes exceed expectations because he or she *"suddenly joins the dots."* Previous work has also shown that teachers with high expectations are less likely than those with low expectations to believe that learning opportunities should be differentiated based on student performance (Rubie, 2003). Together with previous work suggesting that growth mindset teachers generally have higher expectations for students (Rattan et al., 2012), this indicates that fostering a growth mindset in teachers may be one way of reducing unsophisticated task differentiation, but more research is needed to substantiate these initial findings.

Another reason for or against engaging in unsophisticated task differentiation was whether or not teachers considered it to be necessary. Ms Evans and Ms Barns argued that it was not necessary because they already used different types of differentiation, such as differentiation by support. They both also argued that the tasks they provided could at least be accessed by all students even if some would solve fewer and easier tasks. Ms Evans, who had a stronger fixed mindset, explained that differentiation was not necessary because students were grouped by ability and, thus, worked on a similar level. This explanation appears somewhat surprising because Ms Evans taught at Meadows School where students frequently changed ability sets, which seems to show that performance levels varied even within sets. Yet, ability grouping apparently set up her expectations for students, which supports previous findings (Boaler et al., 2000). Ms Evans' justification also provides some support for the proposal that teachers with stronger fixed mindsets may tend to "adjust assignments and teaching between groups but not within groups" (Stipek et al., 2001, p. 223). In contrast, Ms Green, one of the less-fixed mindset teachers, explained that she offered different difficulty levels because students performed at different levels despite being grouped by ability. This suggests that she had a more open view of what ability is because she thought being grouped based on test performance would not characterise students' abilities adequately. Importantly, she did not mandate the difficulty level of tasks herself but instead gave students choice. She believed it was highly important that students were independent learners and took responsibility for their own learning.

Discussions indicated that having performance concerns might explain why teachers use unidimensional tasks, whereas aiming to foster student mastery might be one reason for using multidimensional tasks, in particular tasks that can be solved in different ways (see Section 9.3.1.2). Performance concerns, such as wanting students to do well in exams, seemed to inhibit teachers from teaching alternative strategies. This indicates that these values might be important targets for achieving change in practices, but this should be examined further in future research to substantiate the relationship. In contrast, aiming to foster deep understanding, and thus valuing student mastery, encouraged teachers to provide multidimensional tasks. For instance, the two more-fixed mindset teachers explained that knowing alternative strategies helped students to gain a greater conceptual understanding of the maths work.

Discussions further revealed that using multidimensional tasks, which is an adaptive practice (Boaler, 2016), was not necessarily always justified by adaptive beliefs. Ms Walker argued that letting students choose what methods to use when solving a task was important because she did not want them to struggle. This is indicative of fixed-mindset thinking because it is based on the idea that struggle and mistakes are debilitating for students (Haimovitz & Dweck, 2016, 2017; Sun, 2019). In contrast, Ms Barns, who strongly rejected a fixed mindset, thought that maths was supposed to be difficult, and she did not want to *"dumb students down"* by providing easier methods, although she would always allow students to use different methods given these were correct. Thus, she had high expectations and wanted to challenge students. Considering struggle as motivating and beneficial for learning has previously been argued to reflect a growth mindset (Haimovitz & Dweck, 2017). Importantly, Ms Barns explained that the class collectively searched for strategies, which is in agreement with a multidimensional approach to teaching maths (Sun, 2019), and that she then showed students a worked example that they could follow. The finding that teachers might use adaptive practices for maladaptive reasons indicates that it may not be enough to only target teaching practices without also addressing teachers' beliefs in efforts to foster adaptive student motivation in the long term (e.g., Gregoire, 2003; Yerrick et al., 1997). This will further be discussed in the next chapter (see Section 12.1.5).

11.2.2 Grouping

Although it was initially not of primary interest whether and why teachers provide group and pair work, but rather why they use ability grouping, this emerged as an important theme during interviews. The less-fixed mindset teachers encouraged much more pair work than the more-fixed mindset teachers, but all teachers provided the same explanation for offering pair work: Teachers thought that it was a way of fostering student mastery (see Section 9.3.2.1). According to the TARGET framework, working collaboratively can indeed foster mastery in students (e.g., Ames, 1992a, 1992b; Epstein, 1988). The current findings suggest that teachers who primarily aim to increase students' mastery might be more likely to encourage collaboration amongst students. In contrast, performance concerns of the more-fixed mindset teachers seemed to explain why they offered very little pair work and no group work at all. For instance, Ms Walker did not provide pair work closer to exams, which indicates that she might value test performance over mastery. Additionally, she thought that group work did not provide sufficient structure and clear enough instructions for students to follow. Both more-fixed mindset teachers also said that group work was generally not used at their school because of the *"teaching and learning non-negotiable polic[y],"* which prescribed that everything should be done with a purpose. Ms Evans explained that group work was not considered to be the most direct route to the learning objective, which she seemed to agree with. In general, it was difficult to gauge what the two teachers thought independent of this policy, but it appeared that the underlying assumption was that learning should take the straightest line rather than to be explorative, which points towards a performance rather than mastery emphasis. The findings tentatively indicate that encouraging teachers to focus on students' mastery rather than performance may be one way of promoting their use of collaborative work.

Besides performance concerns, the findings indicate that low expectations might hinder teachers from using group and pair work. For instance, Ms Evans explained that she used pair work with more able students because they did not need to write everything down, which suggests that she might not do this with lower-performing students. Ms Walker thought that there would be no benefit if two students who did not understand something engaged in discussions, and she also did not think that students in Year 7 would *"discover how to do something in math."* Instead, she thought *"knowledge"* needed to be given to students rather than that students developed mathematical ideas together. This reflects a more procedural and unidimensional view of maths, where mostly the teacher does the mathematical work, which has been argued to be more reflective of fixed mindset practices (Sun, 2018, 2019). The current findings suggest that low expectations might be one reason why teachers may focus on following procedures instead of exploration and deep understanding.

Of interest was whether and why teachers group students in heterogeneous or homogenous pairs and, thus, engage in within-class ability grouping. Discussions with teachers did not confirm the expectation that the more-fixed mindset teachers would tend to favour homogenous groups. All teachers but Ms Barns showed a clear preference for heterogeneous pairs (see Section 9.3.2.2). This was surprising because a previous qualitative study found that a fixed mindset teacher preferred homogenous grouping (Rissanen et al., 2018). Yet, Rissanen et al. compared only one fixed mindset teacher with one growth mindset teacher, which is why more work is needed to examine this relationship further. It should also be noted again that the differences between the current teachers' mindsets were not substantial.

All teachers justified forming heterogeneous pairs with the aim to foster mastery and growth, and, unexpectedly, Ms Barns used the same justification for forming homogenous pairs. Teachers thought that students in heterogeneous pairs could help each other and, thus, grow from each other. Ms Green explained though that she would not separate students that worked at the same level as long as they challenged each other and, similarly, Ms Barns argued that both types of pairings could help students improve. It has previously been suggested that forming ability groups within classes and schools is based on notions of a fixed ability, whereas forming heterogeneous groups is more in line with a growth mindset (Boaler, 2013; Marks 2013). Yet, teachers in the current study justified both grouping practices by aiming to foster growth, indicating that beliefs are not necessarily consistent with practices. This may explain why previous research found that even growth mindset teachers group students by ability during lessons (Sun, 2019). This finding suggests that it may be insufficient to target beliefs and practices in isolation. Furthermore, it may be valuable to make teachers aware of potential inconsistencies between their beliefs and practices as well as raise awareness of the potential detrimental consequences of their practices. A meta-analysis of reviews on professional development research found that professional development programmes were most successful if teachers engaged in reflection of practices, such as thinking about the rationales for engaging in a given practice (Cordingley et al., 2015). Potential detrimental effects of FAPs on students' motivational beliefs constitute such a rationale.

Inconsistent with their preference for heterogeneous groups, both more-fixed mindset teachers seated students who struggled with maths towards the front of the class to provide them with more support. This was likely due to low expectations because the teachers seemed to expect that these lower-performing students needed such frequent assistance, that it would become bothersome to walk to them. These teachers seated higher-performing students towards the back of the class, arguing these could work more independently. While it may be accurate that certain students might require more support based on their current ability level, seating these students throughout the classroom should not prevent teachers from providing them with the support they need. Sun (2019) also observed that teachers, even those with growth mindsets, seated weaker students towards the front of the class. She argued that this can be regarded as a type of ability grouping that can send fixed-mindset messages to students. However, more research is needed to confirm this claim. In the current study, a few students also indicated that they thought this practice was beneficial for "weaker" students. Nevertheless, it was interesting that this practice was observed for the more-fixed mindset teachers but not the less fixed-mindset teachers, although one of the latter also seated misbehaving students towards the front to keep an eye on them. This tentatively suggests that teachers' mindsets might play a role in this seating decision, potentially because growth mindset teachers tend to have higher expectations for all students (Rattan et al., 2012). This needs further examination in future research.

Teachers sometimes changed their seating plan, which they justified by aiming to improve students' behaviour and performance (see Section 9.3.2.3). If students performed poorly, Ms Walker tried to pair them with other students in the hope they would improve. Hence, she seemed to attribute students' performance to the situation, in this case with whom they worked and, thus, also to effort because some students would be more focused when seated with a particular student. These adaptive beliefs, in contrast to assuming that students' fixed ability led to poor test performance, seemed to explain the adaptive practice of having flexible pairs. Yet, Ms Walker seemed to only use test performance as an indicator of how successful the seating plan was, which seems like a rather limited view of student success. Ms Barns opposed the practice of changing the seating plan regularly, arguing that it was the teachers' and not the students' responsibility to ensure students are working well together. This seems like a sensible argument, and as long as students are not stuck in homogenous pairs, it might cause no harm to keep the same seating plan throughout the year. Please note that the stability of school-wide ability groups was not discussed with teachers as this is a practice that they had no control over.

11.2.3 Evaluation

All teachers engaged in some form of public evaluation, but the more-fixed mindset teachers did so more frequently and more explicitly. Teachers engaged in public evaluation to motivate students (see Section 9.3.3.1). Ms Walker and Ms Green thought that students would become more motivated and work harder if they saw how well others performed. Ms Walker thought this was particularly true if she praised students that others perceived to be *"dumb."* Similarly, Ms Evans thought that competition created by public evaluation could push students. Previous work showed that having a fixed mindset is positively related to using performance-oriented practices, which included many aspects of public evaluation (Park et al., 2016), and the current study tentatively indicates that this might be due to teachers regarding competition and social comparison as a useful motivational tool. Yet, while competition does not necessarily harm students when it activates beneficial normative-approach goals (e.g., trying to outperform others), it can also foster maladaptive normative-avoidance goals (e.g., trying not to do worse than others; for a review, see Elliot, 2020). Hence, while students may become overall more motivated, teachers seem to risk that the quality of motivation may decline. As discussed further below, student interviews indeed suggested that public evaluation may foster performance goals, in particular the more maladaptive appearance goals.

The findings also indicated that teachers' beliefs were not the only explanation for or against engaging in public evaluation, but that practicalities also mattered and could sometimes even lead teachers to engage in public evaluation despite their disregard for this practice. For instance, even Ms Barns, who thought public evaluation was "cruel," accepted that students could see how others performed for the sake of formative assessment. She and Ms Evans explained they regularly wanted to form momentary snapshots of how students performed, which often entailed public evaluation. Ms Evans also argued that engaging in private evaluation cost too much time. In addition, Ms Walker sometimes wrote the names of students that she planned to award a merit on the board for all to see and even drew verbal attention to this. She argued that she would otherwise forget to follow through. This does not seem particularly plausible as she could have written the names on a piece of paper instead. The question is whether she provided her actual reason or maybe even tried to hide the inconsistency between her behaviour and her previous statement of opposing public evaluation. In any case, addressing such inconsistencies might encourage a deeper reflection of practices.

Besides public evaluation, I identified the theme performance over process emphasis in interviews. Teachers said they focused on effort, behaviour, and improvement rather than performance when evaluating students. Reasons included that teachers valued these aspects, thought that praising these aspects would motivate students, and believed that these were aspects students had more control over (see Section 9.3.3.2). For instance, all teachers but Ms Green explained that they provided process-oriented rather than performance-oriented praise because students had different abilities. Ms Walker also said she encouraged effort because this was what she regarded as the main determinant of performance in lower-ability classes. This is consistent with attribution theory (Weiner et al., 1971), according to which teachers should reinforce attributes that they consider to be responsible for performance. Teachers also focused on the process instead of performance because they valued mastery and thought that maths was not about following a recipe. For example, they said it was important that students conceptually understood what they were doing rather than that they just followed a set of prescribed procedures, which was also why they usually required students to explain how they solved a given task. Notably, Meadows School followed a programme that was specifically aimed at fostering conceptual understanding, which might also partly explain the process emphasis of Ms Evans and Ms Walker. Surprisingly, although Ms Walker said that they aimed to foster conceptual understanding, she tended to display a rather procedural view of maths because she expected students to follow her instructions step by step. Consistent with the cognitive-affective model of conceptual change (Gregoire, 2003) and previous research (e.g., Yerrick et al., 1997), this suggests that simply imposing practices on teachers does not necessarily ensure that teachers adapt their beliefs, which may lead to inconsistencies in the use of practices. This is further discussed later on (see Section 12.1.5).

In general, whether or not teachers used praise frequently seemed to depend on their beliefs about the necessity and potential consequences of praise. Ms Walker, who was the only teacher who praised students very frequently, said that her students became motivated by praise and even expected to be praised for good performance. She thought that immediate praise was necessary because students would otherwise wonder why they were not praised even though they worked well. She did not seem to realise that students' expectations might have only developed because she praised them that often. In contrast, Ms Evans explained that she avoided praising correct answers frequently because this would cause students to feel bad when they gave wrong answers. While the literature suggests that person-oriented praise rather than such product-oriented praise can lead to contingent self-worth (e.g., Kamins & Dweck, 1999), Ms Evans seemed to be aware that praising successes can affect how students feel about failure. Similarly, Ms Green said that she did not use much praise because she regarded it as more important that students learned how to deal with criticism.

11.2.4 Little autonomy

The more-fixed mindset teachers gave students fewer opportunities for choice than the less-fixed mindset teachers. This is consistent with previous research showing that fixed mindset teachers are less likely to establish an autonomy-supportive climate (Leroy et al., 2007). Both teachers justified providing little choice with the importance of structure (see Section 9.3.4.1). They thought structure helped students to understand what was expected of them and, thereby, to work faster, perform better, and according to Ms Walker, also to learn better. Hence, these teachers thought that students were more successful if they provided clear instructions for them to follow, were in control, and made decisions for them. Ms Evans explained that teachers should make decisions for students because of their expertise, educational level, and knowledge of research. She thus prioritised the teacher perspective over the student perspective, which can lead to controlling practices (Reeve, 2009). In addition, teachers at Meadows School had to follow highly-structured lesson plans, which could potentially also explain why they thought structure was so important. Previous work grounded in self-determination theory shows that structure, such as setting clear expectations and rules, are indeed important for supporting students' needs, in particular, their need for competence (see Vansteenkiste et al., 2012). In fact, autonomy support and structure are not considered to be opposites but rather compatible teaching dimensions (see Vansteenkiste et al., 2012). Yet, the two more-fixed mindset teachers seemed to assume that allowing students to make choices and work in their own way were incompatible with the provision of structure. Highlighting this misconception might be one way of encouraging more adaptive teaching practices.

Both less-fixed mindset teachers provided more opportunities for student choice, including choice over tasks and whether they wanted to work in pairs, because they valued student autonomy. Ms Barns revealed implicitly that she valued students as independent thinkers, while Ms Green was very explicit about valuing student responsibility and aiming to train independent learners. Although the more-fixed and less-fixed mindset teachers seemed to differ in their choice provision and valuing of student autonomy, as already suggested by previous research (Stipek et al., 2001), none of them directly related this practice to their mindsets. This might again be due to mindsets being implicit beliefs, which are not as accessible, or even social desirability effects because teachers might have been unwilling to reveal fixed-mindset beliefs. Direct probing might have provided more information about this relationship. Accordingly, more research is needed to examine whether teacher mindset is indeed an antecedent of this practice.

All teachers engaged students in self- or peer-assessment, but only the less-fixed mindset teachers seemed to go beyond that by also engaging students in self-evaluation, including reflecting on areas in need of improvement and monitoring their own progress. These adaptive practices were again explained by adaptive beliefs and values (see Section 9.3.4.2). Ms Walker's and Ms Green's reasons for promoting students' self-regulation included wanting to benefit students' learning and, thus, to foster mastery. In addition, Ms Green explained that self-assessment helped students to take responsibility for their own learning. Interestingly, Ms Barns explained that she wanted her students to engage in self-assessment to gain insights into what they thought about their learning and, thereby, to understand them better. Compared to the more-fixed mindset teachers, Ms Barns thus seemed to have more regard for students' perspective and to welcome their thoughts and feelings, which has been characterised as an autonomy-supportive approach to teaching (see Reeve, 2009).

11.2.5 Differential value

The dimension differential value includes aspects such as paying equal attention to and showing equal interest in all students. However, given that these aspects are highly socially desirable and that teacher discussions would thus have almost certainly been biased, I focused only on the aspect of encouraging participation from all students during interviews.

All four teachers tried to encourage all students, in particular weaker students, to participate during lessons, which they justified by valuing mastery (see Section 9.3.5). Teachers explained they wanted all students to be engaged and learning. While this is clearly a socially desirable stance, so that it seems unlikely that teachers would have said anything contrary, it is consistent with the practices that students reported. Although not said explicitly, teachers' justifications also imply that they assumed that engaging all students, even those that struggled, can be fruitful, which reflects a growth mindset. Similarly, previous qualitative work found that a growth mindset teacher was particularly likely to pay attention to students that struggled or did not participate much (Rissanen et al., 2018). It has been suggested that teachers with stronger fixed mindsets may invest fewer resources in lower-performing students (Leroy et al., 2007; Rissanen et al., 2018). Yet, given that I found no apparent differences between teachers' differential value, there was no support for this suggestion. Considering that the differences in mindsets were small in the current study and that little is known about the relationship between teachers' mindsets and differential value or differential treatment of high- and low-performing students, more work is needed to gain an understanding of such potential links.

Despite the intention to engage all students equally, teachers also acknowledged that they were not always successful in doing so, which indicates that beliefs and intentions do not consistently translate into action. One reason includes that teachers have competing concerns, such as time pressure. Ms Evans explained that she sometimes called on higher-performing students to create pace and quickly move on with the lesson even though she knew that this was not a good practice. Thus, she sometimes sacrificed the understanding of all students for working through all lesson objectives, which points towards pressures created by the school system. Unexpectedly, the less-fixed mindset teachers seemed to accept that some students could not participate in all classroom discussions because certain questions were too difficult. This acceptance appears to be inconsistent with a growth mindset and seems rather problematic because they did not ensure that all students understood their teaching. While it is, of course, important to advance the understanding of high-performers, it might be fairer to challenge these higher-performing students during individual instruction instead of risking that other students become disengaged during whole-class instruction. The interviews could not explain why these less-fixed mindset teachers seemed to accept this discrepancy between their beliefs and practice, but one explanation might be that they were simply not aware of the conflict. Another reason is that they might not know how to ensure that all students can follow their teaching. This suggests that it might be beneficial to raise awareness of these inconsistencies as well as to provide teachers with the skills they need to engage in adaptive practices as will be discussed in the next chapter.

11.2.6 Little promotion of risk-taking

All four teachers reported, and their students largely confirmed, that they encouraged students to ask them for help, which they justified by wanting to ensure that all students understood the learning material and, thus, to foster mastery (see Section 9.3.6.1). Prior research showed that teachers who are themselves oriented towards learning and improving their teaching skills tend to support students' help-seeking (Butler & Shibaz, 2008), but it is unclear why this is the case. Even though teachers in the current study did not refer to their own mastery goals but only to the goals they had for their students, the current findings suggest potential explanations for why this relationship might exist. Specifically, teachers with mastery goals might encourage student help-seeking because they apply the same mastery goal to their students and, thus, want to foster their understanding. This would be consistent with the correspondence hypothesis (i.e., teachers apply their own goals to their students; Dresel et al., 2013). Yet, the functionality hypothesis (i.e., teachers engage in practices that help them reach their own goals; Dresel et al., 2013) also cannot be excluded. For instance, teachers might encourage help-seeking because dealing with students' questions could help them develop their teaching competence. No firm conclusions can be drawn from the current findings, but they suggest a potential explanation for previously identified relationships between teachers' goal orientation and support of help-seeking, which can be explored in future research.

Teachers did not criticise students for making mistakes and even told students that it was OK to make mistakes, which seemed to be explained by their adaptive beliefs and attitudes towards mistakes (see Section 9.3.6.2). They thought everybody made mistakes and that these were valuable for learning, which is more reflective of a growth than a fixed mindset (e.g., Dweck, 1999; Haimovitz & Dweck, 2016). Another explanation may lay in teachers' attributions for mistakes; Ms Green justified not being critical of mistakes with the possibility that she did not explain something adequately and, thus, with the situation students were in. This indicates again that addressing teachers' beliefs may help in efforts to encourage adaptive teaching practices. Additionally, school policies can be influential as Ms Evans and Ms Walker reacted positively to mistakes, amongst other reasons, because it was policy to celebrate mistakes at Meadows School.

Despite teachers' positive attitudes towards mistakes, they did not consistently deal with mistakes in ways that can benefit students' learning, which might partly be due to time and performance pressure. Ms Evans said that she sometimes simply moved on after mistakes to create pace, which indicates that she sacrificed student understanding for working through the entire lesson content. Although not discussed with her, this might be due to performance pressure. Having gotten to know Meadows School and its strong performance emphasis, especially the regular testing regime, it seems plausible that Ms Evans might try to cover as much of the content as possible so that most students performed well on tests, even if that meant a few students were left behind. Sun (2019) also found that even though growth mindset teachers consistently verbalised the importance of mistakes, they did not consistently engage with them in-depth. The present finding tentatively indicates that such inconsistencies might be explained by outside pressures, which suggests that the school environment can impede translating beliefs into practice. This is further discussed in the implications section of the following chapter.

11.2.7 Brief summary on why teachers engage in fixed-ability practices

While all teachers engaged relatively infrequently in FAPs based on their own and student interviews, the more-fixed mindset teachers appeared to engage more frequently in maladaptive practices than the less-fixed mindset teachers. Yet, even the less-fixed mindset teachers did not consistently avoid FAPs. Teachers provided multiple explanations for their practices, including beliefs, values, practical considerations, and school policies. As expected, engagement in more adaptive practices was usually explained by adaptive beliefs or values, whereas maladaptive practices were usually explained by maladaptive beliefs or values as well as practical considerations. Nevertheless, the findings also showed that beliefs and practices were not always consistent. In addition, teachers rarely referred to their mindsets or achievement goals, presumably because these are implicit beliefs and, thus, not as accessible. The contributions and implications of these findings, in particular for teacher professional development, are further discussed in the next chapter (see Sections 12.1.3 and 12.1.5, respectively).

11.3 How do fixed-ability practices affect students' motivational beliefs and achievement?

Overall, the quantitative and qualitative results were consistent with the expectation that student-perceived FAPs foster maladaptive motivational beliefs in students (see Sections 7.2.2 and 10.1, respectively). Yet, some teaching practices seemed to be more influential than others. In particular, students' perceptions of teachers' differential value and little promotion of risk-taking were found to be very harmful, while there were no quantitative effects or discernible trends for student-perceived little promotion of self-regulation. However, student interviews indicated it might potentially strengthen fixed mindsets and inhibit mastery goals. Findings from discussions with students largely aligned with the quantitative findings but they also revealed additional potential links, which might have been too weak to lead to detectable quantitative effects as discussed below. I found no clear evidence that student-perceived FAPs was negatively related to achievement, but there was a quantitative trend suggesting that students' perceiving little promotion of risk-taking performed worse over time. Students' shared perceptions of FAPs at the classroom and teacher level led to an increase in their performance-approach goals. For clarity, this section addresses FAPs one at a time and discusses how these are related to students' motivational beliefs according to the quantitative and qualitative data. Subsequently, I briefly discuss the lack of effects on achievement as well as classroom-level effects.

11.3.1 Tasks

In contrast to expectations, a quantitative trend indicated that students who reported unsophisticated task differentiation might actually become less likely to endorse fixed mindsets over time (see Section 7.2.2.1). While I found no prior empirical studies examining this relationship, it seemed plausible that assigning easier tasks to some students would convey the message that only some students can be successful at solving difficult tasks and, thereby, foster fixed mindsets. One explanation for the unexpected result might be that the scale did not capture the construct as intended. I intended to assess "teacher-determined" task differentiation, while students may have agreed to items even if the teacher let students decide what difficulty level to work on. Interviews suggested that students regarded it negatively if a teacher determined that some students had to work on easier tasks, while they regarded it positively if students could choose their difficulty level (see Section 10.1.1.1).

More specifically, interviews indicated that getting easier tasks might diminish students' striving for mastery as well as their self-concept over time. Students considered unsophisticated task differentiation to be unfair because they thought it reduced students' opportunities for learning. If students regularly received easier tasks than others and were thus not being challenged as much, they may come to believe that improving and understanding the subject matter might be less important than performing well on the tasks they were given. In addition, students reported that getting easier tasks would make them feel as though they were poor *"mathematicians"* and not as intelligent as their peers, which suggests a reduction in their self-concept. Previous research using hypothetical scenarios also found that getting easier tasks following poor performance lowered students' motivation to try to improve as well as their future performance expectations (Rattan et al., 2012). Yet, more research is needed to examine these relationships further because neither that study nor the current interviews provided evidence for causality. The fact that these findings were not supported by the quantitative data might again be explained by questionnaire items having been interpreted differently than intended, which indicates the need for scale refinement.

The unidimensional tasks dimension was not retained in the final questionnaire, as explained above. Hence, only qualitative results are discussed. Contrary to expectations, the interviews did not indicate that engaging in unidimensional tasks, such as tasks that can only be solved in one specific way, might adversely affect students' motivational beliefs (see Section 10.1.1.2). This finding does not confirm the proposal that unidimensional tasks give the impression that maths is a fixed and closed subject, while tasks that can be solved in different ways convey that maths is a "growth subject" (Anderson et al., 2018; Boaler, 2016). Yet, such a relationship cannot be excluded based on the interviews; The leap from solving tasks in alternative ways to motivational beliefs is quite big, so that these connections might not be easily activated without direct probing. In comparison, the link between getting easier tasks and feeling not as capable, as discussed above, seems more straightforward and, thus, more accessible to students.

Interviews indicated that the provision of challenging and rich extension tasks, one aspect of multidimensional tasks, might foster mastery goals over time because students thought these tasks enabled them to push and challenge themselves. Extension tasks also seemed to inhibit performance goals because students reported these tasks prevented them from seeing who finished tasks early since everybody always had something to work on. Students from the relatively performance-oriented Meadows School said that they would not feel bad if they did not reach the extension tasks, which might be due to these tasks being explicitly presented as an additional problem that students were not expected to solve. Interviews thus indicated that there might be potential benefits of providing multidimensional tasks as suggested by Boaler (2016), but more empirical research is needed to examine these effects.

11.3.2 Grouping

The grouping dimensions were not retained in the questionnaire, which is why only qualitative findings can be discussed. Interviews revealed that engaging in group and pair work may convey to students that mastery is important, while not engaging in group work may emphasise the importance of performance (see Section 10.1.2.1). This might, over time, foster corresponding achievement goals. Students appreciated being able to work in groups or pairs because they thought they could exchange ideas and learn from each other. In contrast, students thought teachers may not want them to work together because they should be able to solve problems on their own so that they could do well in tests. Notability, it was not clear whether group work may be related to performance-approach or -avoidance goals, or both because students spoke about the importance of performing well in general terms. Group work has previously been conceptualised as, and shown to be, a practice that fosters mastery goals (Ames, 1992a, 1992b; Lüftenegger et al., 2017). While engagement in group work might thus foster adaptive goal orientations, the current study found no indication that group work may be related to students' mindsets, and I do not know of any studies that investigated this. More research is thus needed to ascertain whether this is a relevant practice to examine when interested in mindsets rather than goal orientations.

Students' opinions of within-class ability grouping or pairing varied considerably (see Section 10.1.2.2). Interviews indicated that homogenous as well as heterogeneous pairing may foster mastery goals over time, depending on which grouping students regard as beneficial for learning. Nevertheless, interviews also suggested that students preferring homogenous pairs were more concerned with performance appearance. Some students preferred homogenous pairs and thought these would increase mastery because similar students would think in similar ways and could receive help from their teacher simultaneously. Others preferred heterogeneous pairs and thought these would benefit mastery because students could help and learn from each other. This could potentially explain why school-wide ability grouping overall does not seem to be associated with students' mastery goals (Butler, 2008). An important question is where these different preferences originate from. Previous exposure to different grouping practices may foster different goal orientations and, thereby, influence future preferences. Indeed, Butler (2008) showed that school-wide homogenous ability grouping fosters performance goals. The current findings suggest that initial goal orientations may influence how students perceive different grouping practices because students with performance appearance concerns seemed to prefer homogenous pairs. Hence, these relationships may be reciprocal in nature, which needs to be tested in future research using longitudinal designs.

Discussions of school-wide ability grouping indicated that ability sets may influence how students perceived their own and their peers' ability and that stable groups may be detrimental to their motivational beliefs (see Sections 10.1.2.2 and 10.1.2.3). Students seemed to accept that those grouped together had a similar ability and, thus, held different expectations for students from different sets. Moreover, they were aware that ability sets can influence their teachers' expectations. These findings are consistent with previous research showing that ability grouping can be related to students' self-concept (Boaler et al., 2000; Reuman, 1989). In line with expectations, discussions also revealed that relatively flexible ability sets may foster growth mindsets because these may communicate to students that they can grow and improve their weaknesses. In contrast, students considered stable ability groups unfair because they provided fewer opportunities to improve. This tentatively supports Boaler's (2013) and Marks' (2013) proposal that stable ability groups might foster fixed mindsets, but much more research is needed to confirm these links. Flexible groupings might also foster mastery goals and effort attributions because some students even reported that moving downwards was fine because this meant they could understand their weaknesses and improve their performance through hard work. However, flexible groupings might also lead to greater performance concerns as other students admitted that they often worried about moving downwards.

Why some students considered moving downwards a threat whilst others thought it was an important learning opportunity was not clear based on interviews. One possibility is that students' initial goal orientations influenced how they perceived the stability of ability sets. Potential reciprocal relationships thus need to be explored further. Butler (2008) proposed that increased performance goals amongst setted students might be due to the salience of social comparison, and current findings suggest that this salience might be exacerbated when students frequently change sets. Accordingly, it would be important to reduce the overall performance emphasis at the school to minimise such performance concerns. Overall, the benefits of flexible sets for students' motivational beliefs seemed to outweigh the potential performance pressure. Whether the motivational processes related to school-wide grouping also apply to within-class grouping, which teachers have more control over, was not answered in this study, which is why more research is needed before recommendations for teachers can be given.

11.3.3 Evaluation

As expected, a quantitative trend indicated that students who thought that their teacher frequently evaluated them in public seemed to adopt more performance-approach goals (see Section 7.2.2.1). Hence, if students noticed that the teacher pointed out their performance and behaviour in front of the class, they likely came to believe that what others thought about their performance and ability mattered. This was supported by student interviews, which indicated potential influences on both performance-approach and -avoidance goals (see Section 10.1.3.1). Students described that some responded to public evaluation with bragging, while others started to shy away from participating or even cheat so that others would not witness their poor performance. This is consistent with many previous studies that found relationships between goal structures or performance-oriented practices and students' goal orientations (e.g., Bardach et al., 2020; Bergsmann et al., 2013). The current study adds to this body of literature by showing that a specific component of performance emphasis, namely public evaluation, may have such detrimental effects.

Interviews also suggested that student-perceived public evaluation may either boost or reduce students' self-concept through social comparison, depending on students' performance. The positive influence on self-concept for some students, such as feeling proud, and the negative one for others, such as feeling humiliated and not as good as others, might explain why there was no quantitative effect of public evaluation on self-concept. The absence of a significant relationship might have masked contrasting patterns of association. Similarly, previous studies found negative or no effects of social comparison and performance goal structures on students' perceived maths ability (Dickhäuser et al., 2017; Givens Rolland, 2012).

It should be noted that there were some students who reported that they did not mind public evaluation, even in the highly performance-oriented classrooms, which might be due to their performance attributions. For instance, one girl described that she was proud of the people who performed well because they had worked hard. This indicates that her attributing success to effort inhibited feelings of competition because she thought people deserved recognition for hard work. This instance highlights the importance of considering motivational beliefs not in isolation but rather as a network of intertwined constructs. Moreover, it suggests that the relationship between FAPs and motivational beliefs is not necessarily a one-way street because the motivational beliefs that students bring into the classroom may influence how they perceive FAPs. It was thus particularly important to control for initial beliefs when examining the effects of FAPs in quantitative analyses.

The dimensions performance over process recognition and social comparison merged into the dimension public evaluation in the quantitative data but a general performance over process emphasis was retained as an important theme in the interviews. It includes aspects such as emphasising tests and speed over mastery as well as giving little process-oriented feedback. Interviews suggested that test and performance emphasis gave students the feeling that they learned in order to perform well rather than to gain mastery, which might lead to the adoption of performance goals and a reduction in mastery goals over time (see Section 10.1.3.2). Similarly, students thought that while speed emphasis might challenge students, it can be humiliating for some students and lead to unnecessary mistakes. These findings are again in line with previous work on performance-oriented practices and students' goal orientation as indicated above. This highlights that teachers may be well advised to reduce performance emphasis in classrooms and, instead, to emphasise the importance of mastery and the learning process. As discussed in the implications section, this would be the responsibility not only of teachers but also of the wider educational system, given that the current accountability measures in the UK may encourage teachers to focus on performance (see Brill et al., 2018).

Interestingly, how students interpreted product-oriented feedback, such as being praised for having solved a problem correctly or having answered many questions, seemed to depend on their attributional style. For instance, some students interpreted getting praised for solving a certain number of tasks as praise for their ability to work fast, while others interpreted it as praise for having focused and worked hard. The former interpretation is more in line with person-oriented praise, which has been found to decrease students' motivation (Haimovitz & Corpus, 2011). Although previous studies did not find maladaptive effects of product-oriented feedback on student motivation (e.g., Corpus & Lepper, 2007), the potentially ambiguous nature of product-oriented feedback suggests that it might be beneficial if teachers linked product-oriented feedback to the process; In interviews, students consistently related process-oriented feedback to a mastery emphasis and the importance of hard work. Previous work already showed that parents' effort praise predicts children's effort attributions (Gunderson et al., 2013), and the current findings tentatively suggest that this may also apply to teaching practices. Yet, more work is needed using longitudinal designs and cross-lagged analyses to examine the reciprocal relationships between teachers' praise and students' attributions.

11.3.4 Little autonomy

The subdimension little choice was not retained in the final student questionnaire for reasons outlined above. Hence, only the interviews can be discussed here. Although most students perceived having little choice rather negatively, thinking their maths work would be easier if they could do things in their own way, the discussions did not reveal any notable links to the motivational beliefs of interest (see Section 10.1.4.1). I expected that having choice would give students a sense of control over their learning, which would foster the belief that they can grow their ability, but this did not become apparent during interviews. This was surprising because previous work found that little opportunities for choice fostered fixed mindsets (Ommundsen, 2001). Findings might differ because Ommundsen's study included items relating to teacher support and promotion of learning in the same factor as student choice, which might have contributed more to the effects on mindsets than the choice components. An interesting question for future research is thus whether autonomy alone is not enough to foster a growth mindset, but that this needs to be accompanied by appropriate teacher guidance, that is the provision of structure. In line with this, research on self-determination theory suggests that a guiding teaching approach that includes both autonomy support and structure is most strongly associated with positive student outcomes (Aelterman et al., 2019). Moreover, the leap from the provision of choice to motivational beliefs seems relatively large; Thinking about having little choice might not easily activate students' thoughts about the motivational beliefs of interest. More direct probing during interviews might have revealed such relationships.

Unexpectedly, there were no quantitative effects indicating potential influences of student-perceived little promotion of self-regulation on students' motivational beliefs or achievement (see Section 7.2.2.1). Yet, discussions with students indicated that there might be some links between the promotion of self-regulation, such as students setting their own goals or evaluating their learning, and a growth mindset as well as mastery goals (see Section 10.1.4.2). Students explained that having to assess or evaluate their work would benefit their learning because it showed them how to improve. They also thought that teachers would be able to adjust their teaching based on what students struggled with. In addition, interviews indicated that students trusted that their teachers would set appropriate goals or targets for them. Hence, students who receive relatively low goals might feel less able over time. Previous studies also found that autonomy support was associated with students' goal orientations (Madjar et al., 2013; Shih, 2013) and self-concept (Deci et al., 1981; Trouilloud et al., 2006). This might not have translated into detectable quantitative effects in the current study because setting learning goals and evaluating their progress may not be part of students' everyday classroom experience, as seen in the classes of the two more-fixed mindset teachers. Instead, it might only be expected of students towards the end of the half-term or after important exams. As such, the practices might not be prevalent enough to make a considerable difference to students' motivation. Hence, although the qualitative data pointed towards potential negative consequences of little promotion of self-regulation, the quantitative results indicate that this dimension might not be as relevant as initially hypothesised, in particular with regard to students' mindsets.

11.3.5 Differential value

Quantitative results showed that students developed stronger fixed mindsets and a lower self-concept when they thought that their teacher valued lower- and higher-performing students differently (see Section 7.2.2.1). This was also supported by the qualitative findings (see Section 10.1.5). Thus, students seemed to infer how able their teachers considered them to be based on how they were being treated compared to others, in particular how often and when they were being called on. This, in turn, seemed to be related to how they viewed their ability. The combination of a lowered self-concept and stronger fixed mindset seems particularly adverse. As far as I know, this is the first study that showed the impact of differential value on students' mindsets, while it substantiates previous research that found negative effects of teachers' differential treatment on students' self-concept (Zhu et al., 2018). Interviews suggested that the adoption of a stronger fixed mindset might be explained by students' assumption that calling mostly high performers indicates that teachers do not believe in other students. When thinking that even their teachers gave up on them, students may come to believe that any efforts to improve will be futile. This is further supported by a quantitative trend which indicated that students began to believe more strongly that failure was determined by a lack of ability, which, in combination with a fixed mindset, suggests that failure is uncontrollable (see Section 7.2.2.1). In contrast, discussions with students indicated that encouraging all students to participate conveys that the teacher believes in them and wants to help them. It has previously been proposed that teachers may invest more resources in high-performing students when they believe that helping low-ability students is futile (Leroy et al., 2007; Rissanen et al., 2018). Even though I did not find that this was the case for teachers in the current study as discussed above, the effects of student-perceived differential value on students' motivation do indicate that students can pick up and internalise this sense of uncontrollability.

Interviews also indicated that encouraging all students to participate may foster mastery goals over time. Students thought that engaging everybody benefitted their learning, whereas only calling on high achievers would not support all students' mastery. Thus, if teachers do not call on struggling students, students may assume that they do not care whether all students understand the learning material, which, over time, might reduce the extent to which students strive for mastery. Yet, the quantitative data did not support this, which indicates that the link might be relatively subtle so that the effects might only become apparent over a longer period. I did not find any previous studies that examined how teachers' differential value affects students' goal orientations. Hence, more research is needed to examine whether this qualitative finding can be substantiated.

11.3.6 Little promotion of risk-taking

A quantitative trend suggested that students who thought their teachers did not promote risk-taking developed a stronger fixed mindset (see Section 7.2.2.1). Consistent with this, students indicated in interviews that constructive handling of mistakes, such as showing students where they went wrong or giving them the chance to try again, conveyed to students the potential for growth, in particular, that they can improve if they invested more effort (see Section 10.1.6.1). This suggests that the finding that parents' debilitating views of mistakes predict children's fixed mindsets (Haimovitz & Dweck, 2016) might also apply to the domain of teaching. Importantly, these results contradict a previous study that found that teachers' encouragement to ask for help predicted stronger fixed mindsets amongst university students (Barger, 2019). One potential explanation is that university students might have already established the assumption that help-seeking is a sign of low ability, whereas the younger students in this study might have been more malleable in this regard. While I could not find studies to support this claim directly, it has been suggested that adults might be less likely to seek help because they focus on the risks of help-seeking, such as being perceived as less competent, while children may focus more on the learning benefits of help-seeking (Nelson-Le Gall, 1985).

Quantitative and qualitative findings suggested that little promotion of risk-taking may be related to students' maladaptive goal orientations (see Sections 7.2.2.1 and 10.1.6, respectively). As expected, quantitative results showed that students developed stronger performance-approach goals by the end of the school year if they thought that their teachers did not promote risk-taking. That is, students began to believe that it was important to demonstrate their smartness in class if they did not feel that teachers treated mistakes as an important part of the learning process and did not carefully answer their questions. Unexpectedly, the quantitative data did not show that these students also adopted stronger performance-avoidance goals despite a positive mean effect. Yet, interviews indicated that both encouragement of help-seeking and constructive handling of mistakes might inhibit the development of performance-avoidance goals because these practices gave students the feeling that appearance concerns were not as important as mastery and improvement. Results are consistent with the findings that fear of failure is an antecedent of performance-approach and -avoidance goals (Elliot & McGregor, 2001) and that students with performance goals tend to avoid help-seeking (Karabenick, 2004). A strong quantitative trend also suggested that students perceiving little promotion of risk-taking were less likely to endorse mastery goals, which was further supported by interviews. For instance, students considered simply moving on after mistakes as useless because students would not learn. If this happened frequently, students might begin to believe that learning was less important than appearing clever. In contrast, students believed that it was important for their teachers to help them learn if they could ask them for help and if teachers answered questions thoroughly. Similarly, discussions indicated that students considered mistakes as important learning opportunities if their teachers gave them the feeling that it was OK to make mistakes. This is consistent with a previous cross-sectional study which found that a positive error climate was positively correlated with students' mastery goals (Steuer et al., 2013).

Interviews also pointed towards the detrimental effects of student-perceived little promotion of risk-taking on students' self-concepts. For instance, one student explained that not being able to ask for help would make him feel dumb, and students revealed that unconstructive handling of mistakes, in particular moving on to other students, would make them feel worse about themselves and less intelligent. This supports the proposal that teachers' judgements of students' ability become apparent when teachers ignore students when they make mistakes (Bohlmann & Weinstein, 2013). Yet, the quantitative data did not show such trends, potentially because any one student might not experience such instances often enough.

11.3.7 Absence of effects on achievement

Contrary to expectations, I did not find clear evidence that students who perceived their teachers to engage in FAPs experienced declines in their maths achievement by the end of the school year (see Section 7.2.2.1). There was only a quantitative trend which suggested that students who thought that their teacher rarely promoted risk-taking experienced declines in achievement. This may be explained by the fact that risk-taking provides greater learning opportunities – when students engage little in class to avoid mistakes and, similarly, if they do not ask questions to avoid looking incompetent, they are presumably less likely to understand the learning content. This is supported by a previous study which showed that the more students asked for help, the more progress they made in maths over one school year (Schenke et al., 2015). More research is needed to substantiate these relationships.

I hypothesised several other effects of the various fixed-ability practices on achievement. For instance, I thought that unsophisticated task differentiation would negatively affect achievement because students who work on easier or fewer tasks would experience less challenge. In addition, there is some prior evidence that performance-oriented practices are negatively related to students' achievement while mastery-oriented practices tend to be positively related to achievement (e.g., Canning et al., 2019; Park et al., 2016; Wolters, 2004). Similarly, teachers' autonomy support has been found to have positive effects on achievement (for a review, see Su & Reeve, 2011), and so does teacher warmth, which is related to differential value, because of increased student participation (Voelkl, 1995). I also expected that achievement might suffer if students are exposed to FAPs because of the negative effects on motivational beliefs, which are, in turn, related to achievement (see Section 2.1 in the literature review). There are several potential explanations for why FAPs did not seem to affect achievement in the current study. One reason is that the FAPs scales were not yet precise enough and need further development as indicated in Section 10.1. Influences of FAPs on achievement may also only become apparent over time. For instance, indirect effects on achievement through motivational beliefs and, thus, learning behaviour may only occur when the beliefs have become more established. This study looked at students in their first year of secondary school, who thus needed to adjust to a completely new school environment. Blackwell et al. (2007) found effects of mindsets on achievement two years after the transition to secondary school, which might be a more reasonable time frame. Note that future work needs to examine whether or not such mediated effects occur. This was not possible in this study because achievement and motivation were assessed at a similar time point.

11.3.8 Classroom level effects

Students' shared perception of FAPs within classrooms and teachers was positively related to performance-approach goals (see Section 7.2.2.2). That is, students in classes that jointly perceived their maths teacher to frequently engage in FAPs increasingly aimed to demonstrate their ability and look smart in front of their teachers and peers. This is not surprising as many FAPs items, including those of the risk-taking dimension as already discussed, also tap into an emphasis on performance (see Table 2). Yet, there were no other effects at the classroom/teacher-level. One potential explanation for the absence of other effects is that there was little variation of the uniform FAPs dimension between teachers and classes. Teachers who frequently engaged in one fixed-ability practice did not necessarily also engage in another fixed-ability practice. Thus, combining all items into one uniform factor likely cancelled differences out. This is indicated by low or even negative correlations between some factors as well as the absence of an overarching FAPs dimension that represented all practices well. The fact that there was an effect on performance-approach goals suggests that these goals were sensitive to even very small variations in the uniform FAPs dimension, much more so than the other motivational beliefs.

It is not unusual that the number of factors for teaching practices at the classroom level is substantially reduced compared to the student level or that the classroom level factor has a smaller effect on students than the individual level factor (e.g., Steuer et al., 2013). Similarly, a factor analysis of mastery goal structures found that student reports could not differentiate between classrooms, while they could differentiate between students within the same classrooms (Lam et al., 2015). This indicates that within-classroom differences can be larger than between-classroom differences. The authors suggested this might be due to the coexistence of different motivational climates within the same classroom or to different student perceptions based on their personality characteristics.

11.3.9 Brief summary on how fixed-ability practices affect students' motivational beliefs and achievement

The quantitative and qualitative results were generally consistent with the expectation that student-perceived FAPs foster maladaptive motivational beliefs in students. Yet, some teaching practices, in particular teachers' differential value and little promotion of risk-taking, seemed to be especially detrimental. Quantitative and qualitative findings were mostly aligned, but the interviews suggested additional links between FAPs and students' motivation, which might have been too subtle to be detected quantitatively. Student-perceived FAPs did not seem to have substantial effects on students' achievement, potentially because the measurements were not yet sharp enough or because these relationships develop over a longer period. Overall, the results supported but also extended the existing literature, which has important implications for various educational stakeholders, as discussed in the next chapter (see Section 12.1).

Chapter 12
Conclusions

After having discussed the findings, this chapter concludes this thesis. It first highlights the contributions and implications of the current study (Section 12.1). It then addresses its limitations and proposes avenues for future research (Section 12.2). Finally, I am making some concluding remarks (Section 12.3).

12.1 Contributions and implications of the study

The current study has made important contributions to the existing knowledge in the field of educational psychology and, in particular, motivation. In this section, I summarise the theoretical, empirical, and methodological contributions, and discuss the practical implications for educators.

12.1.1 Conceptualisation and measurement of fixed-ability practices

One of the key theoretical and methodological contributions of this study are the development of the fixed-ability practices (FAPs) conceptualisation and the corresponding measurement tool for students, respectively. Even though previous frameworks for growth mindset teaching exist (Boaler, 2013; Rissanen et al., 2019; Sun, 2018), this study substantiated and extended these frameworks. For instance, the current study highlighted that differential value, at least as perceived by students, is an important dimension of FAPs and, thus, can direct researchers' attention to this overlooked area. Interestingly, previous frameworks mostly built on relationships between teachers' mindset and their practices rather than on relationships between practices and students' mindsets (Rissanen et al., 2019; Sun, 2018). By focusing on both of these relationships, the current study provided strong grounds for characterising certain teaching practices as *fixed-ability practices*. Importantly, there was not yet a comprehensive measurement tool to assess such practices because prior works on such frameworks have been qualitative in nature. Despite the need for further refinement, particularly of the teacher scales, the developed items are a promising starting point for gaining a fuller picture of FAPs in the classroom by considering multiple dimensions simultaneously. The new student questionnaire thus enables researchers to analyse the impact of distinct but related practices in tandem, which it crucial for understanding the relative contribution of each practice to shaping students' motivational beliefs.

Previous research examining relationships between teaching practices and teachers' and students' mindsets often relied on measures of achievement goal structures or performance-oriented practices (e.g., Dresel et al., 2013; Lüftenegger et al., 2017; Park et al., 2016). As argued in the literature review, while achievement goal orientation is closely related to mindset, these are not identical concepts, which may explain why dimensions such as ability grouping and differential value have not been included in measures focused on goal orientation. Taking a mindset-perspective appears particularly important because student mindset is regarded as a central belief that gives rise to a meaning system consisting of other relevant motivational beliefs, including goal orientation (Payne et al., 2007) but also attribution (Blackwell et al., 2007) and self-concept (Gunderson et al., 2017).

Despite these contributions, the results of the current study as well as studies that have been published after the FAPs conceptualisation has been developed show that more work is needed to refine the conceptualisation and corresponding questionnaire. Most notably, there seem to be theoretical disagreements regarding whether some practices are part of a growth or fixed mindset pedagogy as well as inconsistencies in empirical findings. For instance, assigning easier or less tasks to lower-achieving students can be regarded as a fixed-ability practice as argued in this thesis (also see Boaler, 2016; Rattan et al., 2012; Rissanen et al., 2018), but Rissanen et al. (2019) also argued differentiation to be a growth mindset practice although they did not specify what exactly differentiation constitutes. Similarly, I argued in this thesis that little encouragement of help-seeking can be regarded as a fixed-ability practice (also see Sun, 2019), but one study found that teachers' encouragement of help-seeking fostered university students' fixed mindset (Barger, 2019). As a last example, providing autonomy can be viewed as a growth mindset practice and has been found to be positively related to growth mindsets (see Leroy et al., 2007; Ommundsen, 2001), but I did not find an association between little autonomy support and students' fixed mindset.

I already suggested potential explanations for these inconsistencies in the discussion. Importantly, they highlight the complex nature of fixed-ability practices and the difficulty of categorising a practice as either FAPs or non-FAPs as I set out to do in the literature review. The lines between these two broad practices oftentimes appear blurry. Interviews with teachers and students provided some valuable insights regarding this complexity. They showed that nuances in the actual implementation of practices and the context in which they are being used are crucial for how they are being perceived by students. For example, while I hypothesised that unsophisticated task differentiation would foster students' fixed mindset, there was a statistical trend in the opposite direction. Interviews suggested that this might have been due to a measurement issue. Students perceived it negatively if their teachers determined what difficulty level they should work at, whereas they regarded it as beneficial for learning if they could choose between different difficulty levels themselves. It is thus not simply a matter of "different difficulty/amount for some" vs "same difficulty/amount for all" as suggested in the literature review (see Table 2.1) but also a matter of *how* tasks are being differentiated. These nuances are difficult to capture in questionnaires, which is why more qualitative work is needed to refine the FAPs conceptualisation and questionnaire.

Nevertheless, the new FAPs conceptualisation and student questionnaire can be valuable tools for furthering our knowledge of how fixed mindsets and related maladaptive motivational beliefs develop amongst students. As discussed below, this may benefit efforts to prevent declines in students' motivation not only by targeting these practices directly but also by addressing the reasons why teachers engage in such practices.

12.1.2 Other methodological contributions

Besides developing the student FAPs questionnaire, this study made an important methodological contribution by applying advanced statistical analyses to the field of motivation. Longitudinal designs have distinct advantages over cross-sectional designs when examining causal relationships in an educational setting, but they also bring distinct challenges. One major challenge is that students, teachers, and classrooms are seldom stable groupings. As seen in this study, it is common that classes change teachers, students change classes, or that one class is taught by multiple teachers during a single school year. Traditional multilevel models cannot account for these complex nesting structures, which can lead to a range of statistical biases as discussed (see Section 4.1.2). This study used cross-classified multiple membership models to account for such structures, which enabled me to examine the contribution of practices of multiple teachers weighted by the proportion of time a student was exposed to them. In addition, because some teachers taught multiple classrooms, I was able to distinguish between classroom and teacher effects. This was especially interesting for variations in achievement and self-concept, which showed that although the classroom explained more variance in achievement, most likely due to ability grouping, the teacher explained more variance in self-concept. In sum, applying sophisticated statistical models enabled a more accurate understanding of the effects of teaching practices.

12.1.3 Teachers' motivational beliefs as antecedents of fixed-ability practices

This study made important empirical contributions by exploring why teachers engaged in different fixed-ability practices. This revealed potential antecedents of FAPs, which can provide insight into how such maladaptive practices can be avoided. Consistent with expectations, teacher and student interviews indicated that the fixed-mindset teachers seemed to overall engage in more maladaptive practices, such as allowing little pair work, using frequent public evaluation, and providing little choice. While none of the teachers explicitly justified their practices with their mindsets, this points toward the possibility that fixed mindsets might indeed be related to the use of more FAPs. Overall, teachers' justifications for engaging in a given practice were complex as they did not only include their motivational beliefs and values but also practical considerations and school policies. This has important implications for targeting maladaptive practices as discussed below (see Section 12.1.5).

The current findings agree with the general literature on relationships between teachers' beliefs and their practices which shows that although beliefs can affect practices, these relationships are complex and not always consistent (e.g., Cross, 2009; Pajares, 1992; Raymond, 1997; Sun, 2019; Thompson, 1992; Wilkins, 2008). More specifically, the findings supported previous work, such as that teachers' expectations might be linked to assigning different tasks to students (Rubie, 2003) and that even growth mindset teachers form homogenous ability pairs or groups (Sun, 2019). Moreover, the findings extended previous work by showing that relationships between parents' views of mistakes and their children's mindsets (Haimovitz & Dweck, 2016) may also exist between teachers and their students. The study also provided initial support to ideas that were not yet empirically examined, including that teachers with stronger fixed mindsets may adjust how they teach different ability groups but not how they teach different students within the same ability group (Stipek et al., 2001).

12.1.4 Implications of fixed-ability practices for students' motivational beliefs

By studying the relationships between student-perceived FAPs and students' motivational beliefs, this study generated important empirical findings as discussed in the previous chapter. The results supported previous findings, such as that performance-oriented practices are associated with more maladaptive motivation (e.g., Bardach et al., 2020) and that unsophisticated task differentiation may predict lower self-concept amongst students (Rattan et al., 2012). In addition, it extended previous research by indicating that relationships between parenting practices and children's motivational beliefs can also hold for relationships between teaching practices and students' motivational beliefs, such as that effort praise fosters effort attributions (Gunderson et al., 2013). Importantly, the study also generated new findings including that student-perceived differential value fosters students' fixed mindsets and that little promotion of risk-taking also likely fosters fixed mindsets and did foster performance-approach goals. In addition, the study showed that students' joint perceptions of fixed-ability practices led to an increase in their performance-approach goals. The study also challenged previous work by showing that the finding that teachers' encouragement to ask for help predicted stronger fixed mindsets amongst university students (Barger, 2019) did not apply to a younger age group and can thus not be generalised.

Importantly, while this study was not an experimental one, the longitudinal design gives some confidence that relationships are at least partly causal in nature. This is especially the case because students transitioned to a new school and were thus taught by a new teacher at the beginning of the study. Yet, this does not mean that reciprocal influences do not exist. In fact, the interviews revealed that students' previous motivational beliefs may affect how they perceive teaching practices. Overall, the FAPs framework may help to explain the general decline of motivation after the transition to secondary school (for a review, see Eccles & Roeser, 2009) and can, thereby, also help to identify how this decline may be prevented.

12.1.5 Implications for educational stakeholders

Given the detrimental effects of students' fixed mindsets and related maladaptive motivational beliefs for learning discussed throughout this thesis, it is clear that the development of such beliefs needs to be inhibited. That mindsets can be changed has already been shown twenty years ago in multiple lab studies (see Dweck, 1999). This was then confirmed by multiple intervention studies that focused on students (e.g., Aronson et al., 2002; Blackwell et al., 2007; Good et al., 2003). However, I question whether student interventions are the most feasible and sustainable way of fostering growth mindsets. Even though previous studies were effective, it is unclear how sustained such mindset changes are. Considering that the present study indicated that FAPs, at least as perceived by students, may affect mindsets and associated beliefs, it seems unlikely that students would maintain a growth mindset if they are continuously being exposed to FAPs in their classrooms. This is supported by a study which showed that the initial success of student interventions in fostering growth mindsets does not last if students are then exposed to fixed-ability messages in their learning environment (Schmidt et al., 2015). Accordingly, addressing the antecedents of maladaptive motivational beliefs by targeting the specific teaching practices identified in this study seems like a promising approach. This has important implications for different educational stakeholders. Please note that any practical recommendations arising from this study are made with caution as more research is needed to confirm the current findings.

Foremost, it seems crucial that teachers become aware of the potential harmful consequences of their talk and behaviour which are explicitly or implicitly related to notions of fixed ability. The present findings can help teachers understand what messages regarding the malleability of ability they may convey to students, and how they may unwittingly set up their students to develop maladaptive motivational beliefs. The interviews revealed that all FAPs were perceived mostly negatively by students and that they seemed to adversely affect their motivational beliefs. Yet, student-perceived differential value and little promotion of risk-taking may be particularly damaging as indicated by interviews and the longitudinal data. Hence, even though this research is at its early stages and should be replicated in future studies, educators appear to be well-advised to avoid showing differential value and to encourage student risk-taking. Differential value may be conveyed to students through behaviours such as listening more carefully to high-performing students, asking them to give answers in front of the class more often than low-performing students, as well as showing more interest and giving more attention to high performers. It may thus be beneficial to avoid such behaviours. In contrast, it seems important that teachers give students the feeling that it is ok to take academic risks by handling mistakes constructively, such as by using them as learning opportunities and giving students the chance to try again. Similarly, it appears important that students feel that they can ask for help. It also seems sensible that teachers avoid making public evaluations because a quantitative trend as well as student interviews indicated that this may foster performance-approach goals. Accordingly, it may benefit students if teachers refrained from practices such as discussing and commenting on how well students do in front of the class.

If confirmed by future research, these findings have implications for educational policymakers because initial teacher education and continuing professional development for in-service teachers may provide crucial opportunities for targeting these maladaptive practices. The question is how these teaching practices can be addressed. Discussions with teachers suggest that in order to achieve long-lasting change in practices, it may be necessary to simultaneously a) address underlying motivational beliefs, b) target, and encourage reflection of, specific teaching practices, and c) develop practical skills. Addressing motivational beliefs and values alone in the hope that teachers will naturally begin to avoid maladaptive practices would likely not suffice because teachers' reasons for engaging in practices are diverse and highly complex. While adaptive beliefs and values generally seemed to explain adaptive practices and maladaptive beliefs and values mostly explained maladaptive practices, these relationships were not straightforward. For instance, the more-fixed mindset teachers argued that students should sit in heterogeneous pairs to support each other but they did not actually provide many opportunities for pair discussions. Similarly, the teachers who strongly rejected fixed-mindset beliefs seemed to accept that some students could not engage in difficult classroom discussions although their beliefs should have motivated them to provide explanations that all students could understand. In addition, introducing a practice without addressing beliefs might not be successful. This was demonstrated by one of the more-fixed mindset teachers, who engaged in mostly procedural mathematics although her school advocated student mastery and conceptual understanding over fluency. Lastly, practical skills may need to be fostered besides raising awareness of beliefs and practices because teachers may very well be aware of the potentially damaging effects of certain practices but they may lack the skills to consistently avoid these or feel that they do not have sufficient resources. This is, for instance, supported by the finding that teachers verbally stressed the value of mistakes but did not consistently treat mistakes as learning opportunities.

In addition to these three components, a further important consideration regards the framing of a potential intervention. While this thesis focused on maladaptive fixed-ability practices, it might not be the best strategy to focus only on the "don'ts" when communicating with teachers. Ideally all teachers would be open towards learning about their own "mistakes" in teaching; Yet, when being told to avoid certain teaching practices that they are currently engaging in, teachers may feel criticised and interpret such messages as a threat. Consequently, they may become defensive and reject the intervention rather than engaging deeply with it, in particular if they do not believe they would be able to change their practices (Gregoire, 2003). Moreover, simply telling them that they should avoid certain practices without giving them any guidance for what practices to use instead may leave them feeling helpless. During any intervention, it would thus be important to provide alternative teaching strategies in addition to raising awareness of potentially maladaptive practices. This could include practices such as speaking about the importance of mistakes and helping students to understanding their conceptual underpinnings as well as encouraging all students to actively engage in the lesson. Thereby, one can create an idea of a classroom that teachers can aspire to, which may be more motivating and thus increase teachers' engagement.

All of these implications for teacher training and development are supported by the literature on teachers' beliefs and practices. For instance, introducing practices without addressing beliefs is likely to be unsuccessful because beliefs can act as a filter for any incoming information regarding proposed teaching practices, which can then influence whether or not and to what extent teachers adopt the new practices (e.g., Cross, 2009; Gregoire, 2003; Yerrick et al., 1997). Solely introducing the practice may lead to an integration of that practice into the current belief system without leading to conceptual or practical change (Cross, 2009; Yerrick et al., 1997) or a superficial change in beliefs (Gregoire, 2003). A meta-analysis of reviews on professional development research showed that professional development programmes were most successful if teachers engaged in reflection of the practices, such as thinking about the rationales for engaging in a given practice (Cordingley et al., 2015). The potential detrimental effects of FAPs on students' motivational beliefs constitute such a rationale. Moreover, highlighting the disconnect between certain beliefs and practices might be a powerful way to initiate change because teachers may simply not be aware of the messages their practices can convey. In addition to raising awareness of the potential consequences of FAPs, the underlying motivational beliefs, as well as the potential disconnect between beliefs and practices, teachers then need to learn how to translate theory into actual classroom practices (see Cordingley et al., 2015; Fives & Buehl, 2012; Gregoire, 2003). If teachers do not feel that they can implement new practices, for instance because they lack the necessary skills, knowledge, or resources, they are unlikely to seriously consider adopting them (see Gregoire, 2003).

The current findings also have implications for school leaders because the responsibility for engaging in adaptive teaching practices cannot lay with teachers alone. Teachers do not work in isolation but belong to a specific school context that can either foster or hinder adaptive teaching practices (see Fives & Buehl, 2012). Indeed, interviews suggested that teachers engaged in many practices because their schools' policies encouraged them to do so. For instance, Ms Evans and Ms Walker taught at a highly performance-oriented school that frequently tested students and openly displayed performance tables. In line with this, they engaged in considerably more frequent and explicit public evaluation compared to the teachers at a different school. Their own beliefs and preferences seemed to align with the school policies, but it is plausible that the school policies shaped their beliefs over time. While Ms Evans also showed instances of questioning some school policies, such as giving demerits to students whose test performance declined, it is also the responsibility of her school leaders to ensure that she could engage in adaptive practices. The current study did not examine school policies in detail, but it suggested that this is an area that needs great attention in future research, as suggested in the next section.

Overall, findings of the current study suggest that it may be crucial to address the connections between beliefs and practices in teacher education and interventions because long-lasting positive change is most likely achieved if these work in tandem, especially if the wider school environment is conducive to adaptive practices. I want to emphasise that the FAPs framework is intended to provide a guide for teachers for how they can avoid to unintentionally foster maladaptive beliefs. It is not meant to be prescriptive but to encourage reflections on why certain practices might harm students' motivation. As such, the framework may empower teachers by raising awareness of the effects of certain practices as opposed to taking away their authority and teaching autonomy.

12.2 Limitations and future directions for research

Limitations and suggestions for future research have been made throughout the discussion chapter but there are several key limitations and future research avenues that I want to emphasise here. This section first acknowledges these limitations and points out how they can be approved upon. Subsequently, new directions for future research, based on the results of the study, are considered.

12.2.1 Quantitative and qualitative limitations

Beginning with the quantitative part of the study, the first limitation relates to the perspective from which FAPs were characterised. Initially, I aimed to assess practices from teachers' and students' perspectives. While the student sample was sufficiently large for all analyses, the teacher sample was relatively small, so that the factor structure of FAPs could only be based on student reports. The corresponding teacher FAPs scales were unreliable and could not be used in subsequent analyses. Thus, teachers' practices were only characterised based on students' perspective. The representation of the FAPs construct or its factor structure, and therefore key dimensions, may differ depending on the perspective from which they are being assessed. Practices could have been more wholistically assessed by using observations in addition to student and teacher reports because all three methods have their own strengths and weaknesses and provide access to different elements of the classroom (also see Section 3.2.2 in Chapter 3 Research Methodology). Although maths lessons of four teachers were observed so that their practices could be discussed in teacher and student interviews, these lessons were not coded for FAPs because this went beyond the scope of this PhD. Observations are considered an effective tool for assessing teaching practices, including practices pertaining to classroom goal structures, which are related to FAPs (Wolters et al., 2010). They may also be considered the most direct and objective form of assessment (see Maulana & Helms-Lorenz, 2016; Wolters et al., 2010). Yet, observations also have limitations, including that meaningful behaviours may be missed if they occur infrequently and that nuances in teacher behaviour which shape the overall classroom structures may only develop over time, as well as that teachers are likely to display their best behaviour when being observed (see Maulana & Helms-Lorenz, 2016; Wolters et al., 2010).

Similar to the characterisation of FAPs, the relationship between FAPs and students' motivational beliefs was based on student reports only. The results are consequently susceptible to a common-method bias. As already described, students' previous experiences and characteristics likely influence their perceptions. While I controlled for initial motivation, there may have been other characteristics that could have influenced students' perceptions. Basing recommendations for teaching practices solely on student reports rather than a more objective method may thus be problematic. Yet, student reports have their place in the assessment of teaching practices as described in the methods chapter (see Section 3.2.2), and the adequate reliability of reports within classrooms can give at least some confidence that these were based on actual practices.

Due to the relatively small number of participating teachers, the chance of detecting significant relationships between teachers' motivational beliefs and their engagement in FAPs was very small. I thus also decided not to perform regression analyses to explore non-significant correlations further. The information contained in the teacher data was thus relatively thin and the qualitative data became the main source for understanding why teachers engage in FAPs. A related limitation is that there were only seven schools in the sample, which is why I was unable to model a third level in the analyses to examine between-school variances. A much larger sample of schools may provide additional insights into the variation of FAPs. For instance, if teachers use more similar practices within schools than between schools, there are likely to be school-level factors that influence teachers' use of FAPs, such as common teaching policies regarding the evaluation of students. Teacher interviews already indicated that this is likely to be the case.

Although the results of the present study provided support for the initial validity of the student FAPs questionnaire, additional research is needed to further refine and establish the validity of the instrument as discussed (see Section 11.1). For instance, it needs to be ensured that the unsophisticated task differentiation items are interpreted in the way I intended, and more coherent items assessing unidimensional tasks need to be developed. Moreover, scales assessing ability grouping and choice need to be revised. In addition, I found that the dimensions hold different promise regarding their importance for student motivation. In particular, the dimensions little choice and little promotion of self-regulation might not be as relevant as initially thought. While I do not want to recommend immediately excluding these dimensions, they need to be developed further. It should also be noted again that not all teaching dimensions were represented well by the uniform FAPs factor at the classroom/teacher level, indicating the further need to refine the scales. In future studies, it would also be important to examine the discriminant validity of the scales by comparing it with similar constructs, in particular classroom goal structures, and to test measurement invariance over time because these are relevant for construct validity (see Morin et al., 2013). It would also be interesting to apply the items to other domains than maths, such as English, to examine whether similar practices and motivational processes can be found in other subjects.

While this study made important methodological advances through the application of multiple membership models, students reported practices for only a single teacher, which is why individual perceptions of teaching practices could only be examined for students who were taught by only one maths teacher throughout the year. To take full advantage of methodological developments, future studies should ask students to report on all their maths teachers, which will be feasible since the item number has been drastically reduced in the process of factor analysis.

There are also a few limitations concerning the qualitative part of the study. No teacher with a strong fixed mindset participated in the study, so that the four teachers selected for observations and interviews had only relatively small differences in their mindset. Having a greater range of ability beliefs might have revealed more nuances in FAPs. In addition, it would have been better if these teachers would have taught at different schools. I actually identified a more-fixed mindset teacher from a different school, but he decided not to take part in observations. The fact that teachers with a similar mindset taught at the same school contributed to the difficulty of teasing apart school policy and personal preferences as antecedents of FAPs. Moreover, the interview data was limited by the time that was available for talking to teachers and students. Due to time constraints, some teaching practices were discussed in more detail in some interviews than others, resulting in an imbalance in how much information students and teachers provided for a given practice. Many aspects could have been discussed in more detail, but I was forced to move on to other practices, which limited the data I used to draw conclusions in some instances. More time would also have allowed for more direct probing of teachers' and students' beliefs. I asked teachers why they engaged in certain practices, which was important for exploring their spontaneous justifications. Yet, more time would have allowed to then also address their beliefs more explicitly, which might have provided relevant insight into relationships between beliefs and practices that teachers did not spontaneously reveal themselves, perhaps because beliefs were implicit.

It should also be noted that social desirability might have influenced participants' responses. For instance, one teacher who did not reject fixed-mindset beliefs in questionnaires showed no such beliefs during interviews, while those strongly rejecting fixed-mindset beliefs in questionnaires nevertheless exposed such beliefs during interviews. Social desirability concerns could have also affected how participants acted during lessons. Yet, during interviews, both students and teachers reported that teachers acted mostly as usual, except for teachers being a bit more nervous. Students also reported that they behaved better during the first observed lesson but that their behaviour normalised afterwards.

The findings regarding teachers' beliefs and their practices might not generalise to other contexts given that the overall teacher sample was relatively small and only four teachers participated in the qualitative part of the study, which were not randomly sampled from across the UK or from different year groups (Cohen, 2012). To draw more definite conclusions, future work needs to examine the relationships between teachers' beliefs and practices at a larger scale and in different school contexts. Similarly, this study only included students from Year 7, so that relationships between teaching practices and their motivational beliefs might not generalise to other year groups. Importantly, the schools that participated were rather diverse. Some schools were located in rural and others in urban areas, and their sizes, gender composition, as well as the percentage of students who were eligible for premiums or free school meals varied. Students also came from different ethnic and religious backgrounds. Accordingly, a relatively diverse range of schools and students was considered, which increases the likelihood that similar relationships might exist beyond this specific sample.

12.2.2 New directions for future research

Besides improvements that can be made based on the above limitations, the current findings also provide exciting new avenues for future research. First, it remains important to better understand the relationship between implicit teacher beliefs and practices. This study showed that teachers rarely referred to their own mindset or achievement goals when justifying their practices. It is unclear whether this means that these beliefs are not important antecedents of practices or whether these beliefs are just not as accessible to teachers due to their implicit nature. Probing more directly for such beliefs during interviews and examining these relationships quantitatively with a larger sample may shed more light on this issue. Similarly, more research is needed to understand the relationship between teachers' goal orientations and the goals they have for their students. It has been suggested that teachers may apply the goals they hold for themselves to their students and, thus, engage in practices that foster these goals, but another possibility is that teachers engage in practices that help them reach their own achievement goals (see Dresel et al., 2013). Understanding which of these is the case could inform how adaptive teaching practices can be encouraged.

Second, inconsistencies between teachers' beliefs and their practices need to be examined in greater detail. More qualitative studies are needed to explore together with teachers why they engage in practices that seem inconsistent with their motivational beliefs and values, what they think impedes them from engaging in adaptive practices, and what support and resources they need not only from their schools but also from their initial education and continuing professional development programmes. Current findings suggest that school policies and performance pressures might deter teachers from engaging in more adaptive practices. This is why it is important to continue to examine what notions and practices based on fixed-mindset beliefs may be prevalent at different levels of the educational system. This includes aspects at the school level, such as school policies and the motivational climates that are created by the different values that are being emphasised. It also includes more structural aspects in the educational system such as governments' accountability measures. For instance, in England, national tests are used to hold schools accountable for student performance, which can lead to "teaching to the test" and, thus, a narrowing of the curriculum (see Brill et al., 2018). Aiming to meet performance targets may mean that teachers allocate less time for exploration and, instead, focus on following procedures. The accountability measures can also focus schools' attention and resources towards the students that are just below the threshold of meeting school targets, at the expense of students who might not be perceived as being able to reach the threshold or those that are doing well already (see Brill et al., 2018). As the present study suggested, differential value or investment of fewer resources for certain groups may convey to these disadvantaged students that any efforts to improve may be futile. Future research studying the interplay between these contextual influences and teacher beliefs in influencing teaching practices, and ultimately students' motivational beliefs and achievement, may draw on Bronfenbrenner's ecological systems theory, which shows how the dynamic relationships between different systems drive students' development (Bronfenbrenner, 2009).

Third, it would be interesting to examine whether teachers' reasons for engaging in a given practice make a difference for students' beliefs. Interviews with students showed that students develop ideas about why teaches engage in certain practices, but too little is known about whether their ideas correspond to teachers' justifications. If teachers who engage in the same practice but for different reasons send slightly different messages to students, it would seem especially important to address teachers' underlying beliefs rather than teaching practices alone. Examples include engaging in a practice because teachers are convinced that it is important or because their school mandates it as well as engaging in a practice because teachers want to fulfil their own goals or because they want to fulfil their students' goals.

Fourth, this study indicated that the relationship between student-perceived FAPs and students' motivational beliefs might be reciprocal in nature. While FAPs are likely to shape students' beliefs over time as indicated by the quantitative and qualitative data, discussions with students also suggested that students' initial beliefs may affect how certain practices are perceived. For instance, whether students endorsed mastery or performance goals seemed to affect how they perceived ability grouping and the stability of such groupings. The idea that students' beliefs, personality, and previous experiences can affect how they perceive practices is, of course, not new (e.g., Wolters et al., 2010). Yet, the current findings suggest that disentangling the directionality of effects could shed more light on the interplay between FAPs and motivational beliefs. This could be achieved by conducting a longitudinal study and using cross-lagged analyses in which student-perceived FAPs are not only independent variables but also dependent variables which are predicted by student motivation and achievement.

Fifth, it would be interesting to examine how the constructs of differentiation, that is the adaptation of the learning content, process, and product to individual learners or groups of students, as well as differential treatment of students with high or low ability relate to fixed-ability practices. Rissanen et al. (2019) argued "differentiation becomes the basis of pedagogical practice in a growth mindset pedagogy" (p. 205). Teachers with growth mindsets may be more likely to believe that all students can overcome barriers to learning, such as ineffective learning strategies, and may thus also be more likely to aim to support students' individual learning processes (Rissanen et al., 2019). In contrast, teachers with fixed mindsets may invest less effort into supporting lower achieving students because they may think that helping a student with a stable, low ability is fruitless (Leroy et al., 2007; Rissanen et al., 2018). Differentiating instruction in a way that supports all students is, of course, no easy task, and, if done poorly, may backfire as seen by students' negative perceptions of teachers' *unsophisticated task differentiation* during interviews in the current study. Thus, there seems to be a crucial, but potentially easy to miss, distinction between the kind of sophisticated differentiation or individualisation, where teachers consider students' individual needs to find the optimal way to promote their learning, and the kind of unsophisticated differentiation, where some students simply get easier task and may thus be left behind. One interesting research question would thus be whether teachers with stronger growth mindsets are more likely to engage in "sophisticated" differentiation.

Another interesting research question is whether teachers' mindset is related to their differential treatment of low and high ability students. One related dimension that has already been examined in this thesis is *differential value*. Yet, teachers with a fixed mindset might more generally tend to differentially treat students based on their ability, whereby they might engage in more supportive practices with high ability students. Considering the other FAPs dimensions, it is conceivable that teachers with a fixed mindset might differ in their behaviours related to autonomy support, public evaluation, and reactions to mistakes depending on their perceived ability of students. For example, such a teacher might support the autonomy of high achievers but be controlling with lower achievers. One previous study found that primary school teachers with weaker growth mindsets reported they would use more performance-oriented practices and less mastery-oriented practices with less able students (Miele et al., 2019). Teachers with stronger growth mindsets reported less differential treatment of this kind. Differential treatment of students with high and low perceived ability would have important implications for the assessment of FAPs. In the current study, both teachers and students were asked about the fixed-ability practices that teachers use in their class *in general*. Yet, students within the same class may report different practices because they might be exposed to different practices. These reports would seem less reliable than they actually are, and any differences in reported practices would cancel each other out when aggregating reports to the classroom level. Similarly, teachers may tend to report practices that they use with only certain groups of students or use the middle option of frequency scales. Assessing teaching practices as reports of interactions with individual students in a future iteration of FAPs could thus be interesting. For instance, in Miele et al.'s study, teachers were asked to think about current or former students who had either low or high ability. They then reported how likely they were to engage in different teaching behaviours when said students were struggling with a task. Students could similarly rate how often their teachers engage in certain behaviours when interacting with a low or high performing student.

Lastly, this line of research will eventually require longitudinal interventions to examine how teachers can be encouraged to consistently engage in less FAPs in the long term, and whether this can lead to long-lasting benefits for students' motivation. While the current study showed that student-perceived FAPs may adversely affect students' motivation, more research is needed to substantiate these relationships and to identify the most promising approaches for initiating change in these practices that consider teachers, schools, and the wider educational system.

12.3 Concluding remarks

The longitudinal mixed-methods study reported in this thesis made important contributions to motivational psychology by a) conceptualising teaching practices that are based on the notion that ability is fixed, b) developing a corresponding student questionnaire for the assessment of such fixed-ability practices , c) revealing why teachers' might engage in these maladaptive practices, and d) demonstrating that these practices, at least as perceived by students, may indeed foster students' fixed mindsets and related maladaptive motivational beliefs. The current findings have implications for various educational stakeholders and provide practical guidance for educators regarding what practices may be advisable to avoid in order to weaken the decline of students' motivation. Importantly, the findings emphasise that the responsibility cannot lay with teachers alone, but that the wider educational system needs to support the use of adaptive practices. Overall, this line of research can play an important role in helping students flourish in school.

References

Aelterman, N., Vansteenkiste, M., Haerens, L., Soenens, B., Fontaine, J. R. J., & Reeve, J. (2019). Toward an integrative and fine-grained insight in motivating and demotivating teaching styles: The merits of a circumplex approach. *Journal of Educational Psychology*, *111*(3), 497–521. https://doi.org/10.1037/edu0000293

Ahmavaara, A., & Houston, D. M. (2007). The effects of selective schooling and self-concept on adolescents' academic aspiration: An examination of Dweck's self-theory. *British Journal of Educational Psychology, 77*(3), 613–632. http://dx.doi.org/10.1348/000709906X120132

Ames, C. (1992a). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3)*,* 261–271. https://doi.org/10.1037/0022-0663.84.3.261

Ames, C. (1992b). Achievement goals and the classroom motivational climate. In D. Schunk & J. Meece (Eds.), *Student perceptions in the classroom* (pp. 327–348). Lawrence Erlbaum Associates, Inc.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation process. *Journal of Educational Psychology*, *80*(3), 260–267. https://doi.org/10.1037/0022-0663.80.3.260

Anderman, E. M., Eccles, J. S., Yoon, K. S., Roeser, R., Wigfield, A., & Blumenfeld, P. (2001). Learning to value mathematics and reading: Relations to mastery and performance-oriented instructional practices. *Contemporary Educational Psychology, 26,* 76-95. http://dx.doi.org/ 10.1006/ceps.1999.1043

Anderman, L. H., Patrick, H., Hruda, L. Z., & Linnenbrink, E. A. (2002). Observing classroom goal structures to clarify and expand goal theory. In C. Midgley (Ed.), *Goals, goal structures, and patterns of adaptive learning* (pp. 243–278). Lawrence Erlbaum Associates Publishers.

Anderman, E. M., & Wolters, C. A. (2006). Goals, values, and affect: Influences on student motivation. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 369–389). Lawrence Erlbaum Associates Publishers.

Anderson, R., Boaler, J., & Dieckmann, J. (2018). Achieving elusive teacher change through challenging myths about learning: A blended approach. *Education Sciences*, *8*(3), 98. https://doi.org/10.3390/educsci8030098

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of*

Experimental Social Psychology, 38(2)*,* 113-125. https://doi.org/10.1006/jesp.2001.1491

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 397–438. https://doi.org/10.1080/10705510903008204

Asparouhov, T., & Muthen, B. (2010). *Multiple imputation with Mplus*. Version 2. https://www.statmodel.com/download/Imputations7.pdf

Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: autonomy-enhancing and suppressing teacher behaviors predicting students' engagement in schoolwork. *British Journal of Educational Psychology, 72*(2)*,* 261–278. https://doi.org/10.1348/000709902158883

Attride-Stirling, J. (2001). Thematic networks: An analytic tool for qualitative research. *Qualitative Research*, *1*(3), 385–405. https://doi.org/10.1177/146879410100100307

Babad, E. (1990). Measuring and changing teachers' differential behavior as perceived by students and teachers. *Journal of Educational Psychology, 82*(4), 683–690. https://doi.org/10.1037/0022-0663.82.4.683

Babad, E. (1993). Teachers' differential behavior. *Educational Psychology Review, 5*(4), 347–376. https://doi.org/10.1007/BF01320223

Bahník, Š., & Vranka, M. A. (2017). Growth mindset is not associated with scholastic aptitude in a large sample of university applicants. *Personality and Individual Differences*, *117*, 139–143. https://doi.org/10.1016/j.paid.2017.05.046

Bandura, A. (1997). *Self-efficacy: The exercise of control.* W H Freeman/Times Books/ Henry Holt & Co.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307–337). Information Age Publishing.

Bardach, L., Oczlon, S., Pietschnig, J., & Lüftenegger, M. (2020). Has achievement goal theory been right? A meta-analysis of the relation between goal structures and personal achievement goals. *Journal of Educational Psychology, 112*(6), 1197–1220. https://doi.org/10.1037/edu0000419

Barger, M. M. (2019). Connections between instructor messages and undergraduate students' changing personal theories about education. *The Journal of Experimental Education*, *87*(2), 314–331. https://doi.org/10.1080/00220973.2018.1469111

Bartholomew, K. J., Ntoumanis, N., Mouratidis, A., Katartzi, E., Thøgersen-Ntoumani, C., & Vlachopoulos, S. (2018). Beware of your teaching style: A school-year long investigation of controlling teaching and student motivational experiences. *Learning and Instruction*, *53*, 50–63. https://doi.org/10.1016/j.learninstruc.2017.07.006

Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and random effects models: Making an informed choice. *Quality & Quantity*, *53*(2), 1051–1074. https://doi.org/10.1007/s11135-018-0802-x

Bennett, D. A. (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health*, *25*(5), 464–469. https://doi.org/10.1111/j.1467-842X.2001.tb00294.x

Beretvas, S. N. (2011). Cross-classified and multiple-membership models. In J. J. Hox & J. K. Roberts (Eds.), *European Association for Methodology series. Handbook for advanced multilevel analysis* (pp. 313–334). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9780203848852.ch16

Berger, J. O., & Bayarri, M. J. (2004). The interplay of Bayesian and Frequentist analysis. *Statistical Science*, *19*(1), 58–80. https://doi.org/10.1214/088342304000000116

Bergsmann, E. M., Lüftenegger, M., Jöstl, G., Schober, B., & Spiel, C. (2013). The role of classroom structure in fostering students' school functioning: A comprehensive and application-oriented approach. *Learning and Individual Differences*, *26*, 131–138. https://doi.org/10.1016/j.lindif.2013.05.005

Biddle, C., & Schafft, K. A. (2015). Axiology and anomaly in the practice of mixed methods work: Pragmatism, valuation, and the transformative paradigm. *Journal of Mixed Methods Research*, *9*(4), 320–334. https://doi.org/10.1177/1558689814533157

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246–263. https://doi.org/10.1111/j.1467-8624.2007.00995.x

Blatchford, P., Hallam, S., Ireson, J., Kutnick, P., & Creech, A. (2008). Classes, groups and transitions: structures for teaching and learning. *Primary Review Research Survey, 9*(2), 548–588. University of Cambridge Faculty of Education.

Boaler, J. (2013). Ability and mathematics: The mindset revolution that is reshaping education. In *Forum* (Vol. 55, pp. 143–152). Symposium Journals. http://www.youcubed.org/wp-content/uploads/14_Boaler_FORUM_55_1_web.pdf

Boaler, J. (2016). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching.* Jossey-Bass.

Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record,* 110(3), 608–645.

Boaler, J., Wiliam, D., & Brown, M. (2000). Students' experiences of ability grouping-disaffection, polarisation and the construction of failure. *British Educational Research Journal*, *26*(5), 631–648. https://doi.org/10.1080/713651583

Bohlmann, N. L., & Weinstein, R. S. (2013). Classroom context, teacher expectations, and cognitive level: Predicting children's math ability judgments. *Journal of Applied Developmental Psychology*, *34*(6), 288–298. https://doi.org/10.1016/j.appdev.2013.06.003

Bong, M. (2004). Academic motivation in self-efficacy, task value, achievement goal orientations, and attributional beliefs. *The Journal of Educational Research*, *97*(6), 287–298. https://doi.org/10.3200/JOER.97.6.287-298

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1–40. https://doi.org/10.1023/A:1021302408382

Bostwick, K. C. P., Collie, R. J., Martin, A. J., & Durksen, T. L. (2020). Teacher, classroom, and student growth orientation in mathematics: A multilevel examination of growth goals, growth mindset, engagement, and achievement. *Teaching and Teacher Education*, *94*, 103100. https://doi.org/10.1016/j.tate.2020.103100

Bostwick, K. C. P., Martin, A. J., Collie, R. J., & Durksen, T. L. (2019). Growth orientation predicts gains in middle and high school students' mathematics outcomes over time. *Contemporary Educational Psychology*, *58*, 213–227. https://doi.org/10.1016/j.cedpsych.2019.03.010

Boyatzis, R. E. (1998). *Transforming qualitative information*. Sage Publications, Inc.

Bråten, I., & Strømsø, H. I. (2006). Predicting achievement goals in two different academic contexts: A longitudinal study. *Scandinavian Journal of Educational Research, 50*(2), 127–148. https://doi.org/10.1080/00313830600575932

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2019). Thematic Analysis. In P. Liamputtong (ed.), *Handbook of research methods in health social sciences* (pp. 843–860). Springer

Brill, F. J., Grayson, L., Kuhn, L., & O'Donnell, S. (2018). *What impact does accountability have on curriculum, standards and engagement in education? A literature review*. NFER

British Educational Research Association (2011). *Ethical guidelines for educational research.* British Educational Research Association.

http://www.bera.ac.uk/files/2011/08/BERA-Ethical-Guidelines-2011.pdf.
British Psychological Society (2014). *Code of human research ethics*. BPS. http://www.bps.org.uk/system/files/Public%20files/inf180_web.pdf.
Bronfenbrenner, U. (2009). *The ecology of human development*. Harvard University Press.
Brophy, J. E., & Good, T. L. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology, 61*(5), 365–374. https://doi.org/10.1037/h0029908
Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005
Browne, W. J., Goldstein, H., & Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. *Statistical Modelling, 1,* 103–124. https://doi.org/10.1177%2F1471082X0100100202
Bürkner, P.-C. (2017). Brms: An *R* package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01
Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017
Bürkner, P.- C., & Charpentier, E. (2020). Modeling monotonic effects of ordinal predictors in regression models. *British Journal of Mathematical and Statistical Psychology.* https://doi.org/10.1111/bmsp.12195
Bürkner, P.-C., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science*, *2*(1), 77–101. https://doi.org/10.1177/2515245918823199
Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, *139*(3), 655–701. https://doi.org/10.1037/a0029531
Butler, R. (2000). Making judgments about ability: The role of implicit theories of ability in moderating inferences from temporal and social comparison information. *Journal of Personality and Social Psychology*, *78*(5), 965–978. https://doi.org/10.1037/0022-3514.78.5.965
Butler, R. (2006). An achievement goal perspective on student help seeking and teacher help giving in the classroom: Theory, research, and educational implications. In S. Karabenick & R. Newman (Eds.), *Help seeking in academic settings: Goals, groups, and contexts* (pp. 15–44). Lawrence Erlbaum Associates.
Butler, R. (2007). Teachers' achievement goal orientations and associations with teachers' help seeking: Examination of a novel approach to teacher motivation. *Journal of*

Educational Psychology, *99*(2), 241–252. https://doi.org/10.1037/0022-0663.99.2.241

Butler, R. (2008). Ego-involving and frame of reference effects of tracking on elementary school students' motivational orientations and help seeking in math class. *Social Psychology of Education*, *11*(1), 5–23. https://doi.org/10.1007/s11218-007-9032-0

Butler, R. (2012). Striving to connect: Extending an achievement goal approach to teacher motivation to include relational goals for teaching. *Journal of Educational Psychology*, *104*(3), 726–742. https://doi.org/10.1037/a0028613

Butler, R., & Shibaz, L. (2008). Achievement goals for teaching as predictors of students' perceptions of instructional practices and students' help seeking and cheating. *Learning and Instruction*, *18*(5), 453–467. https://doi.org/10.1016/j.learninstruc.2008.06.004

Canning, E. A., Muenks, K., Green, D. J., & Murphy, M. C. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Science Advances*, *5*(2). https://doi.org/10.1126/sciadv.aau4734

Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, *42*(12), 1150–1152. https://doi.org/10.1111/j.1365-2923.2008.03172.x

Carpenter B., Gelman A., Hoffman M. D., Lee D., Goodrich B., Betancourt M., Brubaker M., Guo J., Li P., & Riddell A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*. 76(1). https://doi.org/10.18637/jss.v076.i01

Charter, R. A. (2003). Study samples are too small to produce sufficiently precise reliability coefficients. *The Journal of General Psychology*, *130*(2), 117–129. https://doi.org/10.1080/00221300309601280

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Chen, J. A., & Tutwiler, M. S. (2017). Implicit theories of ability and self-efficacy: Testing alternative social cognitive models to science motivation. *Zeitschrift Für Psychologie*, *225*(2), 127–136. https://doi.org/10.1027/2151-2604/a000289

Cheng, Z., & Hau, K.-T. (2003). Are intelligence and personality changeable? Generality of Chinese students' beliefs across various personal attributes and age groups. *Personality and Individual Differences*, *34*(5), 731–748. https://doi.org/10.1016/S0191-8869(02)00030-2

Chestnut, E. K., Lei, R. F., Leslie, S.-J., & Cimpian, A. (2018). The myth that only brilliant people are good at math and its implications for diversity. *Education Sciences*, *8*(2), 65. https://doi.org/10.3390/educsci8020065

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Church, M. A., Elliot, A. J., & Gable, S. L. (2001). Perceptions of classroom environment, achievement goals, and achievement outcomes. *Journal of Educational Psychology, 93*(1), 43–54. https://doi.org/10.1037/0022-0663.93.1.43

Cohen, L. (2012). *Research methods in education* (7th ed.) Routledge

Cooper, H. M. (1979). Pygmalion grows up: A model for teacher expectation communication and performance influence. *Review of Educational Research*, *49*(3), 389–410. https://doi.org/10.3102/00346543049003389

Cooper, H., & Good, T. (1983). *Pygmalion grows up: Studies in the expectation communication process.* Longman.

Cordingley, P., Higgins, S., Greany, T., Buckler, N., Coles-Jordan, D., Crisp, B., Saunders, L., Coe, R. (2015). *Developing great teaching: Lessons from the international reviews into effective professional development. Project report.* Teacher Development Trust, London. http://dro.dur.ac.uk/15834/1/15834.pdf?DDD45+DDD29+DDO128+hsmz78+d700tmt

Corpus, J. H., & Lepper, M. R. (2007). The effects of person versus performance praise on children's motivation: Gender and age as moderating factors. *Educational Psychology*, *27*(4), 487–508. https://doi.org/10.1080/01443410601159852

Costa, A., & Faria, L. (2018). Implicit theories of intelligence and academic achievement: A meta-analytic review. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00829

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, & Evaluation, 10*, 1-9. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.110.9154&rep=rep1&type=pdf

Crabtree, B. & Miller, W. (1999). A template approach to text analysis: Developing and using codebooks. In B. Crabtree & W. Miller (Eds.), *Doing qualitative research* (pp. 163–177). Sage Publications, Inc.

Creswell, J. W., & Plano-Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage Publications, Inc.

Cross, D. I. (2009). Alignment, cohesion, and change: Examining mathematics teachers' belief structures and their influence on instructional practices. *Journal of Mathematics Teacher Education*, *12*(5), 325–346. https://doi.org/10.1007/s10857-009-9120-5

Cury, F., Elliot, A. J., Da Fonseca, D., & Moller, A. C. (2006). The social-cognitive model of achievement motivation and the 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology, 90*(4), 666–679. https://doi.org/10.1037/0022-3514.90.4.666

Dai, T., & Cromley, J. G. (2014). Changes in implicit theories of ability in biology and dropout from STEM majors: A latent growth curve approach. *Contemporary Educational Psychology*, *39*(3), 233–247. https://doi.org/10.1016/j.cedpsych.2014.06.003

De Castella, K. & Byrne, D. (2015). My intelligence may be more malleable than yours: the revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. *European Journal of Psychology and Education, 30,* 245–267. https://doi.org/10.1007/s10212-015-0244-y

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology, 73*(5), 642–650. https://doi.org/10.1037/0022-0663.73.5.642

Deemer, S. (2004). Classroom goal orientation in high school classrooms: Revealing links between teacher beliefs and classroom environments. *Educational Research*, *46*(1), 73-90. https://doi.org/10.1080/0013188042000178836

Depaoli, S., & Van de Schoot, R. (2017). Improving transparency and replication in Bayesian statistics: The WAMBS-Checklist. *Psychological Methods*, *22*(2), 240–261. https://doi.org/10.1037/met0000065

DeSimone, J. A., DeSimone, A. J., Harms, P. D., & Wood, D. (2018). The differential impacts of two forms of insufficient effort responding. *Applied Psychology*, *67*(2), 309–338. https://doi.org/10.1111/apps.12117

DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical Education*, *40*(4), 314–321. https://doi.org/10.1111/j.1365-2929.2006.02418.x

Dickhäuser, O., Janke, S., Praetorius, A.-K., & Dresel, M. (2017). The effects of teachers' reference norm orientations on students' implicit theories and academic self-concepts.

Zeitschrift Für Pädagogische Psychologie, *31*, 205–219. https://doi.org/10.1024/1010-0652/a000208

Diseth, Å., Meland, E., & Breidablik, H. J. (2014). Self-beliefs among students: Grade level and gender differences in self-esteem, self-efficacy and implicit theories of intelligence. *Learning and Individual Differences, 35,* 1–8. https://doi.org/10.1016/j.lindif.2014.06.003

Dresel, M., Fasching, M. S., Steuer, G., Nitsche, S., & Dickhauser, O. (2013). Relations between teachers' goal orientations, their instructional practices and students' motivation. *Psychology, 4*(7), 572–584. https://doi.org/10.4236/psych.2013.47083

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development.* Psychology Press.

Dweck, C. (2015). Carol Dweck revisits the "growth mindset". Education Week. https://www.edweek.org/ew/articles/2015/09/23/carol-dweck-revisits-the-growth-mindset.html?cmp=eml-eb-subcontent-mindset+10292017

Dweck, C. (2016, December 16). How praise became a consolation prize [Interview with C. Gross-Loh]. The Atlantic. Retrieved from https://www.theatlantic.com/education/archive/2016/12/how-praise-became-a-consolation-prize/510845/

Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, *6*(4), 267–285. https://doi.org/10.1207/s15327965pli0604_1

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256–273. https://doi.org/10.1037/0033-295X.95.2.256

Dweck, C. S., & Molden, D. C. (2017). Mindsets: Their impact on competence motivation and acquisition. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation* (2nd ed., pp. 135–154). Guilford Press.

Eccles, J. S., O'Neill, S. A. & Wigfield, A. (2005). Ability self-perceptions and subjective task values in adolescents and children. In K. A. Moore & L. H. Lippman (Eds.), *What do children need to flourish?* (Vol. 3, pp. 237–249). Springer.

Eccles, J. S., & Roeser, R. W. (2009). Schools, academic motivation, and stage-environment fit. In R. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (pp. 404–434). Wiley.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the achiever: The structure of adolescents' academic achievement-related beliefs and self-perceptions. *Personality and Social*

Psychology Bulletin, 21, 215–225.

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*(3), 830–847. https://doi.org/10.2307/1131221

Education Endowment Foundation (2018). *Setting or streaming*. https://educationendowmentfoundation.org.uk/pdf/generate/?u=https://educationendowmentfoundation.org.uk/pdf/toolkit/?id=127&t=Teaching%20and%20Learning%20Toolkit&e=127&s=

Elliot, A. J. (2005). A conceptual history of the achievement goal construct. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (Vol. 16, pp. 52-72). Guilford Press.

Elliot, A. J. (2020). Competition and achievement outcomes: A hierarchical motivational analysis. *Motivation Science, 6*(1), 3–11. https://doi.org/10.1037/mot0000164

Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*(3), 501–519. https://doi.org/10.1037/0022-3514.80.3.501

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Epstein, J. L. (1988). Effective schools or effective students: Dealing with diversity. In R. Haskins & D. MacRae (Eds.), *Policies for America's public schools: Teacher, equity and indicators* (pp. 89–126). Ablex.

Fang, Z. (1996). A review of research on teacher beliefs and practices. *Educational Research*, *38*(1), 47–65. https://doi.org/10.1080/0013188960380104

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics*, *21*(1), 55–69.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, *5*(1), 80–92. https://doi.org/10.1177/160940690600500107

Field, A. P. (2009). *Discovering statistics using SPSS* (3rd ed). Sage Publications, Inc.

Fielding, A., & Goldstein, H. (2006). *Cross-classified and multiple membership structures in multilevel models: An introduction* and review, Department of Education and Skills, Birmingham. Available at:

https://webarchive.nationalarchives.gov.uk/20130321021729/https:/www.education.gov.uk/publications/eOrderingDownload/RR791.pdf

Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors* (pp. 471–499). American Psychological Association. https://doi.org/10.1037/13274-019

Gabry, J., & Mahr, T. (2020). Bayesplot: plotting for Bayesian models. R package version 1.7.2. https://cran.r-project.org/web/packages/bayesplot/index.html

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *182*(2), 389–402. https://doi.org/10.1111/rssa.12378

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis (*3rd ed.) Taylor and Francis.

Gelman, A., & Hill, J., (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, *2*(4), 1360–1383. https://doi.org/10.1214/08-AOAS191

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7, 457–511. https://doi.org/10.1214/ss/1177011136

Gelman, A., Simpson, D., & Betancourt, M. (2017). The prior can generally only be understood in the context of the likelihood. *Entropy*, *19*(10), 555. https://doi.org/10.3390/e19100555

George, S. V., & Richardson, P. W. (2019). Teachers' goal orientations as predictors of their self-reported classroom behaviours: An achievement goal theoretical perspective. *International Journal of Educational Research*, *98*, 345–355. https://doi.org/10.1016/j.ijer.2019.09.011

Georgiou, S. N., Christou, C., Stavrinides, P., & Panaoura, G. (2002). Teacher attributions of student failure and teacher behavior toward the failing student. *Psychology in the Schools*, *39*(5), 583–595. https://doi.org/10.1002/pits.10049

Gillham, B. (2000). *The research interview.* Continuum

Givens Rolland, R. (2012). Synthesizing the evidence on classroom goal structures in middle and secondary schools: A meta-analysis and narrative review. *Review of Educational Research*, *82*, 396-435. https://doi.org/10.3102%2F0034654312464909

Goldstein, H. (2003). Multilevel modelling of educational data. In D. Courgeau (Ed.), *Methodology and epistemology of multilevel analysis* (Vol. 2, pp. 25–42). Springer. https://doi.org/10.1007/978-1-4020-4675-9_2

Goldstein, H., Burgess, S., & McConnell, B. (2007). Modelling the effect of pupil mobility on school differences in educational achievement. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *170*(4), 941–954. https://doi.org/10.1111/j.1467-985X.2007.00491.x

Gomm, R. (2004). *Social research methodology. A critical introduction.* Palgrave Macmillan.

Gonida, E., Kiosseoglou, G., & Leondari, A. (2006). Implicit theories of intelligence, perceived academic competence, and school achievement: Testing alternative models. *The American Journal of Psychology, 119*(2), 223–238. https://doi.org/10.2307/20445336

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, *24*(6), 645–662. https://doi.org/10.1016/j.appdev.2003.09.002

Good, T. L., & Nichols, S. L. (2001). Expectancy effects in the classroom: A special focus on improving the reading performance of minority students in first-grade classrooms. *Educational Psychologist*, *36*(2), 113–126. https://doi.org/10.1207/S15326985EP3602_6

Gottschall, A. C., West, S. G., & Enders, C. K. (2012). A comparison of item-level and scale-level multiple imputation for questionnaire batteries. *Multivariate Behavioral Research*, *47*(1), 1–25. https://doi.org/10.1080/00273171.2012.640589

Gravetter, F. & Forzano, L-A. (2012). *Research methods for the behavioral sciences*. Wadsworth Cengage Learning.

Greany, T., Barnes, I., Mostafa, T., Pensiero, N., & Swensson, C. (2016). *Trends in Maths and Science Study (TIMSS): National Report for England*. http://dera.ioe.ac.uk/28040/1/TIMSS_2015_England_Report_FINAL_for_govuk_-_reformatted.pdf

Gregoire, M. (2003). Is it a challenge or a threat? A dual-process model of teachers' cognition and appraisal processes during conceptual change. *Educational Psychology Review, 15*(2), 147–179. https://doi.org/10.1023/A:1023477131081

Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin- Meadow, S., & Levine, S. C. (2013). Parent praise to 1- to 3-year-olds predicts children's motivational frameworks 5 years later. *Child Development, 84,* 1526–1541. http://dx.doi.org/10.1111/cdev.12064

Gunderson, E. A., Hamdan, N., Sorhagen, N. S., & D'Esterre, A. P. (2017). Who needs innate ability to succeed in math and literacy? Academic-domain-specific theories of intelligence about peers versus adults. *Developmental Psychology*, *53*(6), 1188–1205. https://doi.org/10.1037/dev0000282

Hackel, T. S., Jones, M. H., Carbonneau, K. J., & Mueller, C. E. (2016). Re-examining achievement goal instrumentation: Convergent validity of AGQ and PALS. *Contemporary Educational Psychology*, *46*, 73–80. https://doi.org/10.1016/j.cedpsych.2016.04.005

Haimovitz, K., & Corpus, J. H. (2011). Effects of person versus process praise on student motivation: stability and change in emerging adulthood. *Educational Psychology, 31*(5), 595–609. https://doi.org/10.1080/01443410.2011.585950

Haimovitz, K., & Dweck, C. S. (2016). What predicts children's fixed and growth intelligence mind-sets? Not their parents' views of intelligence but their parents' views of failure. *Psychological Science, 27*(6), 859–869. https://doi.org/10.1177/0956797616639727

Haimovitz, K. & Dweck, C.S. (2017). The origins of children's growth and fixed mindsets: New research and a new proposal. *Child Development, 88*, 1849–1859. https://doi.org/10.1111/cdev.12955

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Pearson.

Hallam, S., Ireson, J., Lister, V., Chaudhury, I., & Davies, J. (2003). Ability grouping in the primary school: A survey. *Educational Studies, 29* (1), 69–83. https://doi.org/10.1080/03055690303268

Harackiewicz, J. M., Barron, K. E., Pintrich, P. R., Elliot, A. J., & Thrash, T. M. (2002). Revision of achievement goal theory: Necessary and illuminating. *Journal of Educational Psychology, 94*(3), 638–645. https://doi.org/10.1037/0022-0663.94.3.638

Harris, D. M. (2012). Varying teacher expectations and standards: curriculum differentiation in the age of standards-based reform. *Education and Urban Society*, *44*(2), 128–150. https://doi.org/10.1177/0013124511431568

Heisig, J. P., Schaeffer, M., & Giesecke, J. (2017). The costs of simplicity: Why multilevel models may benefit from accounting for cross-cluster differences in the effects of controls. *American Sociological Review*, *82*(4), 796–827.

https://doi.org/10.1177/0003122417717901

Higgins, S., Katsipataki, M., Kokotsaki, D., Coleman, R., Major, L. E., & Coe, R. (2013). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit*. Retrieved from http://dro.dur.ac.uk/11453/1/11453.pdf

Hong, Y.-y., Chiu, C.-y., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology, 77*(3), 588–599. https://doi.org/10.1037/0022-3514.77.3.588

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53–60. https://doi.org/10.21427/D7CF7R

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin*, *136*(3), 422–449. https://doi.org/10.1037/a0018947

IBM Corp. (2017). IBM SPSS Statistics for Mac (Version 25.0) [Computer software]. IBM Corp.

Ireson, J., & Hallam, S. (2001). *Ability grouping in education.* Sage Publications, Inc.

Ireson, J., Hallam, S., Hack, S., Clark, H., & Plewis, I. (2002). Ability grouping in English secondary schools: Effects on attainment in English, mathematics and science. *Educational Research and Evaluation*, *8*(3), 299–318. https://doi.org/10.1076/edre.8.3.299.3854

Jacobs, N. W., Berduszek, R. J., Dijkstra, P. U., & van der Sluis, C. K. (2017). Validity and reliability of the upper extremity work demands scale. *Journal of Occupational Rehabilitation*, *27*(4), 520–529. https://doi.org/10.1007/s10926-016-9683-9

Jacobson, E., & Izsák, A. (2015). Knowledge and motivation as mediators in mathematics teaching practice: The case of drawn models for fraction arithmetic. *Journal of Mathematics Teacher Education*, *18*(5), 467–488. https://doi.org/10.1007/s10857-015-9320-0

Jang, H.-R. (2017). Teachers' intrinsic vs. extrinsic instructional goals predict their classroom motivating styles. *Learning and Instruction, 60*, 286–300. https://doi.org/10.1016/j.learninstruc.2017.11.001

Joffe, H. (2012). Thematic analysis. In D. Harper & A.R. Thompson (Eds.), *Qualitative research methods in mental health and psychotherapy: A guide for students and practitioners* (pp. 209-223). John Wiley & Sons, Ltd.

Jonsson, A.-C., Beach, D., Korp, H., & Erlandson, P. (2012). Teachers' implicit theories of intelligence: influences from different disciplines and scientific theories. *European Journal of Teacher Education*, *35*(4), 387–400. https://doi.org/10.1080/02619768.2012.662636

Jussim, L., & Eccles, J. S. (1992). Teacher expectations: II. Construction and reflection of student achievement. *Journal of Personality and Social Psychology, 63*(6), 947–961. https://doi.org/10.1037/0022-3514.63.6.947

Käfer, J., Kuger, S., Klieme, E., & Kunter, M. (2019). The significance of dealing with mistakes for student achievement and motivation: Results of doubly latent multilevel analyses. *European Journal of Psychology of Education*, *34*(4), 731–753. https://doi.org/10.1007/s10212-018-0408-7

Kagan, D. M. (1992). Implication of research on teacher belief. *Educational Psychologist*, *27*(1), 65–90. https://doi.org/10.1207/s15326985ep2701_6

Kamins, M. L., & Dweck, C. S. (1999). Person versus process praise and criticism: Implications for contingent self-worth and coping. *Developmental Psychology, 35*(3), 835–847. https://doi.org/10.1037/0012-1649.35.3.835

Karabenick, S. A. (1998). *Strategic help seeking: Implications for learning and teaching*. Lawrence Erlbaum Associates.

Karabenick, S. A. (2004). Perceived achievement goal structure and college student help seeking. *Journal of Educational Psychology*, *96*, 569–581. https://doi.org/10.1037/0022-0663.96.3.569

Kelley, K. (2017). MBESS (Version 4.0.0 and higher) [computer software and manual]. https://CRAN.R-project.org/package=MBESS

King, R. B. (2019). Mindsets are contagious: The social contagion of implicit theories of intelligence among classmates. *British Journal of Educational Psychology, 90*(2), 349–363. https://doi.org/10.1111/bjep.12285

Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design.* Methuen.

Kuklinski, M. R., & Weinstein, R. S. (2001). Classroom and developmental differences in a path model of teacher expectancy effects. *Child Development, 72*(5), 1554–1578. https://doi.org/10.1111/1467-8624.00365

Kutnick, P., Blatchford, P. & Baines, E. (2005). Grouping of pupils in secondary school classrooms: Possible links between pedagogy and learning. *Social Psychology of Education, 8*, 349–374. https://doi.org/10.1007/s11218-005-1212-1

Lam, A. C., Ruzek, E. A., Schenke, K., Conley, A. M., & Karabenick, S. A. (2015). Student perceptions of classroom achievement goal structure: Is it appropriate to aggregate? *Journal of Educational Psychology*, *107*(4), 1102–1115. https://doi.org/10.1037/edu0000028

Lambert, P. C., Sutton, A. J., Burton, P. R., Abrams, K. R., & Jones, D. R. (2005). How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Statistics in Medicine*, *24*(15), 2401–2428. https://doi.org/10.1002/sim.2112

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Lee, K. (1996). A study of teacher responses based on their conceptions of intelligence. *The Journal of Classroom Interaction*, *31*(2), 1–12.

Lee, M., & Bong, M. (2016). In their own words: Reasons underlying the achievement striving of students in schools. *Journal of Educational Psychology*, *108*(2), 274–294. https://doi.org/10.1037/edu0000048

Leondari, A., & Gialamas, V. (2002). Implicit theories, goal orientations, and perceived competence: impact on students' achievement behaviour. *Psychology in the Schools, 39*(3)*,* 279–291. https://doi.org/10.1002/pits.10035

Leroy, N., Bressoux, P., Sarrazin, P., & Trouilloud, D. (2007). Impact of teachers' implicit theories and perceived pressures on the establishment of an autonomy supportive climate. *European Journal of Psychology of Education, 22*(4), 529–545. https://doi.org/10.1007/BF03173470

Liu, Y., Millsap, R. E., West, S. G., Tein, J.-Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological Methods*, *22*(3), 486–506. https://doi.org/10.1037/met0000075

Lüdtke, O., & Köller, O. (2002). Individuelle Bezugsnormorientierung und soziale Vergleiche im Mathematikunterricht: Einfluss unterschiedlicher Referenzrahmen auf das fachspezifische Selbstkonzept der Begabung [Individual reference norm and social comparisons in mathematics classes: The impact of different frames of reference on the domain-specific self-concept of ability]. *Zeitschrift für Entwicklungspsychologie* *und Pädagogische Psychologie, 34*(3), 156–166. https://doi.org/10.1026//0049-8637.34.3.156

Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology*, *34*(2), 120–131. https://doi.org/10.1016/j.cedpsych.2008.12.001

Lüftenegger, M., & Chen, J. A. (2017). Implicit theories: International perspectives from multiple domains. *Zeitschrift Für Psychologie*, *225*(2), 97–98. https://doi.org/10.1027/2151-2604/a000285

Lüftenegger, M., Kollmayer, M., Bergsmann, E., Jöstl, G., Spiel, C., & Schober, B. (2015). Mathematically gifted students and high achievement: The role of motivation and classroom structure. *High Ability Studies*, *26*(2), 227–243. https://doi.org/10.1080/13598139.2015.1095075

Lüftenegger, M., Tran, U. S., Bardach, L., Schober, B., & Spiel, C. (2017). Measuring a mastery goal structure using the TARGET framework: Development and validation of a classroom goal structure questionnaire. *Zeitschrift Für Psychologie*, *225*, 64–75. https://doi.org/10.1027/2151-2604/a000277

Lüftenegger, M., van de Schoot, R., Schober, B., Finsterwald, M., & Spiel, C. (2014). Promotion of students' mastery goal orientations: Does TARGET work? *Educational Psychology*, *34*(4), 451–469. https://doi.org/10.1080/01443410.2013.814189

Luo, W., Lee, K., Ng, P., & Ong, J. (2014). Incremental beliefs of ability, achievement emotions and learning of Singapore students. *Educational Psychology, 34,* 619–634. https://doi.org/10.1080/01443410.2014.909008

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1*(3), 86–92. https://doi.org/10.1027/1614-2241.1.3.86

Madjar, N., Nave, A., & Hen, S. (2013). Are teachers' psychological control, autonomy support and autonomy suppression associated with students' goals? *Educational Studies*, *39*(1), 43–55. https://doi.org/10.1080/03055698.2012.667871

Marks, R. (2013). "The blue table means you don't have a clue": the persistence of fixed-ability thinking and practices in primary mathematics in English schools. *Forum: For Promoting 3-19 Comprehensive Education*, *55*(1), 31–44. http://doi.org/10.2304/forum.2013.55.1.31

Marsh, H. W., Ellis, L. A., Parada, R. H., Richards, G., & Heubeck, B. G. (2005). A short version of the Self Description Questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment, 17*(1), 81–102. https://doi.org/10.1037/1040-3590.17.1.81

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for tit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*(3), 471–491. https://doi.org/10.1037/a0019227

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*(1), 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 439–476. https://doi.org/10.1080/10705510903008220

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology, 111*(2), 331–353. https://doi.org/10.1037/edu0000281

Marshall, H. H., & Weinstein, R. S. (1984). Classroom factors affecting students' self-evaluations: An interactional model. *Review of Educational Research, 54*(3), 301–325. https://doi.org/10.2307/1170451

Mascret, N., Elliot, A. J., & Cury, F. (2017). The 3 × 2 Achievement Goal Questionnaire for Teachers. *Educational Psychology*, *37*(3), 346–361. https://doi.org/10.1080/01443410.2015.1096324

Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research*, *3*(1), 97-110. https://doi.org/10.21500/20112084.854

Matthes, B., & Stoeger, H. (2018). Influence of parents' implicit theories about ability on parents' learning-related behaviors, children's implicit theories, and children's academic achievement. *Contemporary Educational Psychology*, *54*, 271–280. https://doi.org/10.1016/j.cedpsych.2018.07.001

Maulana, R., & Helms-Lorenz, M. (2016). Observations and student perceptions of the quality of preservice teachers' teaching behaviour: Construct representation and predictive quality. *Learning Environments Research*, *19*(3), 335–357. https://doi.org/10.1007/s10984-016-9215-8

McCutchen, K. L., Jones, M. H., Carbonneau, K. J., & Mueller, C. E. (2016). Mindset and standardized testing over time. *Learning and Individual Differences*, *45*, 208–213. https://doi.org/10.1016/j.lindif.2015.11.027

McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan.* CRC Press.

McIntosh, C. N. (2007). Rethinking fit assessment in structural equation modelling: A commentary and elaboration on Barrett (2007). *Personality and Individual Differences*, *42*(5), 859–867. https://doi.org/10.1016/j.paid.2006.09.020

Mears, C. (2009). *Interviewing for education and social science research.* Palgrave Macmillan Ltd.

Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology, 57,* 487–503. https://doi.org/10.1146/annurev.psych.56.091103.070258

Merriam, S. B. & Tisdell, E. J. (2015). *Qualitative research.* (4th ed.) JosseyBass.

Meyers, J. L., & Beretvas, S. N. (2006). The impact of inappropriate modeling of cross-classified data structures. *Multivariate Behavioral Research*, *41*(4), 473–497. https://doi.org/10.1207/s15327906mbr4104_3

Midgley, C., Anderman, E., & Hicks, L. (1995). Differences between elementary and middle school teachers and students: A goal theory approach. *The Journal of Early Adolescence, 15*(1), 90–113. https://doi.org/10.1177/0272431695015001006

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., Gheen, K., Kaplan, A., Kumar, R., Middleton, M. J., Nelson, J., Roeser, R., & Urdan, T. (2000). *Manual for the Patterns of Adaptive Learning Scales*. Ann Arbor, MI: University of Michigan. http://www.umich.edu/~pals/PALS%202000_V12Word97.pdf

Miele, D. B., Perez, S. A., Butler, R., Browman, A. S., O'Dwyer, L. M., & McNeish, D. (2019, August 27). Elementary school teachers' growth mindsets predict their differential treatment of high versus low ability students. *PsyArXiv*. https://doi.org/10.31234/osf.io/qcd83

Miles, M. B., Huberman, A. M., & Saldaäna, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications, Inc.

Miller, A. D., & Murdock, T. B. (2007). Modeling latent true scores to determine the utility of aggregate student perceptions as classroom indicators in HLM: The case of classroom goal structures. *Contemporary Educational Psychology*, *32*(1), 83–104. https://doi.org/10.1016/j.cedpsych.2006.10.006

Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Review of Educational Research*, *90(3), 376–419*. https://doi.org/10.3102/0034654320919354

Moorman, E. A., & Pomerantz, E. M. (2010). Ability mindsets influence the quality of mothers' involvement in children's learning: An experimental investigation. *Developmental Psychology*, *46*(5), 1354–1362. https://doi.org/10.1037/a0020376

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(1), 116–139. https://doi.org/10.1080/10705511.2014.961800

Morin, A. J., Marsh, H. W., & Nagengast, B. (2013). Exploratory structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 395–436). Information Age Publishing, Inc.

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology, 75*(1), 33–52. https://doi.org/10.1037/0022-3514.75.1.33

Muenks, K., Miele, D. B., Ramani, G. B., Stapleton, L. M., & Rowe, M. L. (2015). Parental beliefs about the fixedness of ability. *Journal of Applied Developmental Psychology*, *41*, 78–89. https://doi.org/10.1016/j.appdev.2015.08.002

Muir, T. (2010). Using video-stimulated recall as a tool for reflecting on the teaching of mathematics. In L. Sparrow, B. Kissane, & C. Hurst (Eds.), *Shaping the future of mathematics education: Proceedings of the 33rd annual conference of the Mathematics Education Research Group of Australasia.* MERGA. http://eric.ed.gov/?id=ED520951

Muthén, B., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, *38*(2), 171–189. https://doi.org/10.1111/j.2044-8317.1985.tb00832.x

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide* (8th ed.). Muthén & Muthén.

Natale, K., Aunola, K., & Nurmi, J.-E. (2009). Children's school performance and their parents' causal attributions to ability and effort: A longitudinal study. *Journal of Applied Developmental Psychology*, *30*(1), 14–22. https://doi.org/10.1016/j.appdev.2008.10.002

Nelson-Le Gall, S. (1985). Help-seeking behavior in learning. *Review of Research in Education*, *12*, 55–90. https://doi.org/10.2307/1167146

Neuendorf, K. A. (2019). Content analysis and thematic analysis. In P. Brough (Ed.), *Research methods for applied psychologists: Design, analysis and reporting* (pp. 211-223). Routledge.

Nolen, S. B. (2003). Learning environment, motivation, and achievement in high school science. *Journal of Research in Science Teaching, 40*(4)*,* 347–368. https://doi.org/10.1002/tea.10080

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*(1), 1–13. https://doi.org/10.1177/1609406917733847

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, *19*, 1–13. https://doi.org/10.1177/1609406919899220

OECD (2019). *PISA 2018 Results (Volume I): What Students Know and Can Do*. OECD. https://doi.org/10.1787/5f07c754-en

Ommundsen, Y. (2001). Students' implicit theories of ability in physical education classes: The influence of motivational aspects of the learning environment. *Learning Environments Research*, *4*, 139–158. https://doi.org/10.1023/A:1012495615828

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*(3), 307–332. https://doi.org/10.3102/00346543062003307

Park, D., Gunderson, E. A., Tsukayama, E., Levine, S. C., & Beilock, S. L. (2016). Young children's motivational frameworks and math achievement: Relation to teacher-reported instructional practices, but not teacher theory of intelligence. *Journal of Educational Psychology*, *108*(3), 300–313. https://doi.org/10.1037/edu0000064

Park, D., Yu, A., Baelen, R. N., Tsukayama, E., & Duckworth, A. L. (2018). Fostering grit: Perceived school goal-structure predicts growth in grit and grades. *Contemporary Educational Psychology*, *55*, 120–128. https://doi.org/10.1016/j.cedpsych.2018.09.007

Patrick, H., Anderman, L. H., Ryan, A. M., Edelin, K. C., & Midgley, C. (2001). Teachers' communication of goal orientations in four fifth-grade classrooms. *The Elementary School Journal*, *102*(1), 35–58. https://doi.org/10.1086/499692

Patton, M. Q. (2002). Enhancing the quality and credibility of qualitative analysis. In M. Q. Patton (Ed.), *Qualitative research and evaluation methods,* (pp. 541-598). Thousand Oaks.

Payne, S. C., Youngcourt, S. S., & Beaubien, J. M. (2007). A meta-analytic examination of the goal orientation nomological net. *Journal of Applied Psychology, 92*(1), 128–150. https://doi.org/10.1037/0021-9010.92.1.128

Pintrich, P. R. (1990). Implications of psychological research on student learning and college teaching for teacher education. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 826-857). Macmillan.

Pintrich, P. R. (2000). An achievement goal theory perspective on issues in motivation terminology, theory, and research. *Contemporary Educational Psychology*, *25*(1), 92–104. https://doi.org/10.1006/ceps.1999.1017

Pomerantz, E. M., & Kempner, S. G. (2013). Mothers' daily person and process praise: Implications for children's theory of intelligence and motivation. *Developmental Psychology*, *49*(11), 2040–2046. https://doi.org/10.1037/a0031840

Pomerantz, E. M., & Saxon, J. L. (2001). Conceptions of ability as stable and self-evaluative processes: A longitudinal examination. *Child Development, 72*(1), 152–173. https://doi.org/10.1111/1467-8624.00271

Powell, E. (2005). Conceptualising and facilitating active learning: Teachers' video-stimulated reflective dialogues. *Reflective Practice*, *6*(3), 407–418. https://doi.org/10.1080/14623940500220202

QSR International Pty Ltd (2018). NVivo qualitative data analysis software (Version 12) [Computer software].

R Core Team (2017) *R: A language and environment for statistical computing.* https://www.R-project.org/

Ramirez, G., Hooper, S. Y., Kersting, N. B., Ferguson, R., & Yeager, D. (2018). Teacher math anxiety relates to adolescent students' math achievement. *AERA Open*, *4*(1), 1–13. https://doi.org/10.1177/2332858418756052

Rasbash, J., & Browne, W. J. (2008). *Non-hierarchical multilevel models.* In J. de Leeuw & E. Meijer (Eds.), *Handbook of multilevel analysis* (pp. 301–334). Springer Science + Business Media. https://doi.org/10.1007/978-0-387-73186-5_8

Rattan, A., Good, C., & Dweck, C. S. (2012). "It's ok — Not everyone can be good at math":

Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology, 48*(3), 731–737. https://doi.org/10.1016/j.jesp.2011.12.012

Raubenheimer, J. (2004). An item selection procedure to maximise scale reliability and validity. *SA Journal of Industrial Psychology*, *30*(4), 59–64. https://doi.org/10.4102/sajip.v30i4.168

Raudenbush, S. W. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher*, *34*(5), 25–31. https://doi.org/10.3102/0013189X034005025

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage Publications, Inc.

Raymond, A. M. (1997). Inconsistency between a beginning elementary school teacher's mathematics beliefs and teaching practice. *Journal for Research in Mathematics Education*, *28*(5), 550–576. https://doi.org/10.2307/749691

Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, *44*(3), 159–175. https://doi.org/10.1080/00461520903028990

Retelsdorf, J., Butler, R., Streblow, L., & Schiefele, U. (2010). Teachers' goal orientations for teaching: Associations with instructional practices, interest in teaching, and burnout. *Learning and Instruction*, *20*(1), 30–46. https://doi.org/10.1016/j.learninstruc.2009.01.001

Retelsdorf, J., & Günther, C. (2011). Achievement goals for teaching and teachers' reference norms: Relations with instructional practices. *Teaching and Teacher Education*, *27*(7), 1111–1119. https://doi.org/10.1016/j.tate.2011.05.007

Reuman, D. A. (1989). How social comparison mediates the relation between ability-grouping practices and students' achievement expectancies in mathematics. *Journal of Educational Psychology, 81*(2), 178–189. https://doi.org/10.1037/0022-0663.81.2.178

Rissanen, I., Kuusisto, E., Hanhimäki, E., & Tirri, K. (2018). Teachers' implicit meaning systems and their implications for pedagogical thinking and practice: A case study from Finland. *Scandinavian Journal of Educational Research*, *62*(4), 487–500. https://doi.org/10.1080/00313831.2016.1258667

Rissanen, I., Kuusisto, E., Tuominen, M., & Tirri, K. (2019). In search of a growth mindset pedagogy: A case study of one teacher's classroom practices in a Finnish elementary school. *Teaching and Teacher Education*, *77*, 204–213. https://doi.org/10.1016/j.tate.2018.10.002

Robins, R. W., & Pals, J. L. (2002). Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self and Identity*, *1*(4), 313–336. https://doi.org/10.1080/15298860290106805

Rubie, C. M. (2003). *Expecting the best: Instructional practices, teacher beliefs and student outcomes*. https://researchspace.auckland.ac.nz/handle/2292/28

Rubie-Davies, C. (2009). Teacher expectations and labeling. In L. J. Saha & A. G. Dworkin (Eds.), *International handbook of research on teachers and teaching* (pp. 695–707). Springer US. https://doi.org/10.1007/978-0-387-73317-3_43

Ruzek, E. A., & Schenke, K. (2019). The tenuous link between classroom perceptions and motivation: A within-person longitudinal study. *Journal of Educational Psychology, 111*(5), 903–917. https://doi.org/10.1037/edu0000323

Sainani, K. L. (2015). Dealing with missing data. *PM&R*, *7*(9), 990–994. https://doi.org/10.1016/j.pmrj.2015.07.011

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment, 29*(4), 347–363. http://dx.doi.org/10.1177/0734282911406661

Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(2), 167–180. https://doi.org/10.1080/10705511.2014.882658

Schad, D. J., Betancourt, M., & Vasishth, S. (2019). *Toward a principled Bayesian workflow in cognitive science. ArXiv:1904.12765 [Stat]*. http://arxiv.org/abs/1904.12765

Schafer, J.L. (1999) Multiple imputation: A primer. *Statistical Methods in Medical Research, 8*, 3–15. http://dx.doi.org/10.1191/096228099671525676.

Schenke, K., Lam, A. C., Conley, A. M., & Karabenick, S. A. (2015). Adolescents' help seeking in mathematics classrooms: Relations between achievement and perceived classroom environmental influences over one school year. *Contemporary Educational Psychology*, *41*, 133–146. https://doi.org/10.1016/j.cedpsych.2015.01.003

Schiefele, U., & Schaffner, E. (2015). Teacher interests, mastery goals, and self-efficacy as predictors of instructional practices and student motivation. *Contemporary Educational Psychology*, *42*, 159–171. https://doi.org/10.1016/j.cedpsych.2015.06.005

Schmidt, J. A., Shumow, L., & Kackar-Cam, H. (2015). Exploring teacher effects for mindset intervention outcomes in seventh-grade science classes. *Middle Grades Research Journal, 10*(2), 17–32.

Schwinger, M., Steinmayr, R., & Spinath, B. (2016). Achievement goal profiles in elementary school: Antecedents, consequences, and longitudinal trajectories. *Contemporary Educational Psychology*, *46*, 164–179. https://doi.org/10.1016/j.cedpsych.2016.05.006

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, *77*(4), 454–499. https://doi.org/10.3102/0034654307310317

Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (3rd ed). Teachers College Press.

Senko, C., & Dawson, B. (2017). Performance-approach goal effects depend on how they are defined: Meta-analytic evidence from multiple educational outcomes. *Journal of Educational Psychology*, *109*(4), 574–598. https://doi.org/10.1037/edu0000160

Senko, C., Durik, A. M., & Harackiewicz, J. M. (2008). Historical perspectives and new directions in achievement goal theory: Understanding the effects of mastery and performance-approach goals. In J. Y. Shah & W. L. Gardner (Eds.), *Handbook of motivation science* (p. 100–113). The Guilford Press.

Sewasew, D., & Schroeders, U. (2019). The developmental interplay of academic self-concept and achievement within and across domains among primary school students. *Contemporary Educational Psychology*, *58*, 204–212. https://doi.org/10.1016/j.cedpsych.2019.03.009

Shanley, L., Biancarosa, G., Clarke, B., & Goode, J. (2019). Relations between mathematics achievement growth and the development of mathematics self-concept in elementary and middle grades. *Contemporary Educational Psychology*, *59*. https://doi.org/10.1016/j.cedpsych.2019.101804

Shih, S.-S. (2013). Autonomy support versus psychological control, perfectionism, and Taiwanese adolescents' achievement goals. *The Journal of Educational Research*, *106*(4), 269-279. https://doi.org/10.1080/00220671.2012.692734

Shim, S. S., Cho, Y., & Cassady, J. (2013). Goal structures: The role of teachers' achievement goals and theories of intelligence. *Journal of Experimental Education, 81*(1), 84–104. https://doi.org/10.1080/00220973.2011.635168

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., & Macnamara, B. N. (2018). To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analyses. *Psychological Science*, *29*(4), 549–571. https://doi.org/10.1177/0956797617739704

Skaalvik, E. M., & Skaalvik, S. (2013). School goal structure: Associations with students? perceptions of their teachers as emotionally supportive, academic self-concept, intrinsic motivation, effort, and help seeking behavior. *International Journal of Educational Research*, *61*, 5-14. https://doi.org/10.1016/j.ijer.2013.03.007

Skipper, Y., & Douglas, K. (2012). Is no praise good praise? Effects of positive feedback on children's and university students' responses to subsequent failures: Effects of positive feedback. *British Journal of Educational Psychology*, *82*(2), 327–339. https://doi.org/10.1111/j.2044-8279.2011.02028.x

Snijders, T. A. (2005). Fixed and random effects. *Encyclopedia of Statistics in Behavioral Science*, *2*(2), 664–665. https://doi.org/10.1002/0470013192.bsa234

Spruce, R., & Bol, L. (2015). Teacher beliefs, knowledge, and practice of self-regulated learning. *Metacognition and Learning*, *10*(2), 245–277. https://doi.org/10.1007/s11409-014-9124-0

Stan Development Team (2018a). *RStan: the R interface to Stan*. R package version 2.17.3. https://mc-stan.org/rstan

Stan Development Team (2018b). *ShinyStan: Interactive visual and numerical diagnostics and posterior analysis for Bayesian models*. R package version 2.5.0. http://mc-stan.org

Standholtz, J. H., Ogawa, R. T., & Scribner, S. P. (2004). Standards gaps: Unintended consequences of local standards-based reform. *Teachers College Record, 106*(6), 1177-1202. http://dx.doi.org/10.1111/j.1467-9620.2004.00376.x

Steuer, G., Rosentritt-Brunn, G., & Dresel, M. (2013). Dealing with errors in mathematics classrooms: Structure and relevance of perceived error climate. *Contemporary Educational Psychology*, *38*(3), 196–210. https://doi.org/10.1016/j.cedpsych.2013.03.002

Stevenson, H., Lee, S.-Y., Chen, C., Stigler, J., Hsu, C., Kitamura, S., & Hatano, G. (1990). Contexts of achievement: A study of American, Chinese, and Japanese Children. *Monographs of the Society of Research in Child Development*, *55*(1-2), 1-119. https://doi.org/10.2307/1166090

Stipek, D. J., Givvin, K. B., Salmon, J. M., & MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and Teacher Education*, *17*(2), 213–226. https://doi.org/10.1016/S0742-051X(00)00052-4

Stipek, D., Salmon, J. M., Givvin, K. B., Kazemi, E., Saxe, G., & MacGyvers, V. L. (1998). The value (And convergence) of practices suggested by motivation research and promoted by mathematics education reformers. *Journal for Research in Mathematics Education*, *29*(4), 465–488. https://doi.org/10.2307/749862

Su, Y. L. & Reeve, J. (2011). A meta-analysis of the effectiveness of intervention programs designed to support autonomy. *Educational Psychology Review, 23,* 159–188. https://doi.org/10.1007/s10648-010-9142-7

Sun, K. L. (2018). The role of mathematics teaching in fostering student growth mindset. *Journal for Research in Mathematics Education, 49*(3), 330–355. https://doi.org/10.5951/jresematheduc.49.3.0330

Sun, K. L. (2019). The mindset disconnect in mathematics teaching: A qualitative analysis of classroom instruction. *The Journal of Mathematical Behavior*, *56*. https://doi.org/10.1016/j.jmathb.2019.04.005

Suri, H. (2011). Purposeful sampling in qualitative research synthesis. *Qualitative Research Journal*, *11*(2), 63–75. https://doi.org/10.3316/QRJ1102063

Tabachnik, B. (2014). *Using multivariate statistics.* Pearson Education.

Tanner, H. & Jones, S. (2007). Learning from children about their learning with and without ICT using video-stimulated reflective dialogue. In J. Watson & K. Beswick (Eds.), *Essential research: Essential practice* (Proceedings of the 30th annual conference of the Mathematical Education Research Group of Australasia, pp. 708-716). MERGA.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences* (Vol. 1). Sage.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, *1*(1), 77–100. https://doi.org/10.1177/1558689806292430

Tendeiro, J. N., & Kiers, H. A. L. (2019). A review of issues about null hypothesis Bayesian testing. *Psychological Methods, 24*(6), 774–795. https://doi.org/10.1037/met0000221

Thompson, A. (1992). Teachers' beliefs and conceptions: A synthesis of research. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127–146). Macmillan.

Trouilloud, D., Sarrazin, P., Bressoux, P., & Bois, J. (2006). Relation between teachers' early expectations and students' later perceived competence in physical education classes: Autonomy-supportive climate as a moderator. *Journal of Educational Psychology*, *98*(1), 75–86. https://doi.org/10.1037/0022-0663.98.1.75

Tulis, M. (2013). Error management behavior in classrooms: Teachers' responses to student mistakes. *Teaching and Teacher Education*, *33*, 56–68. https://doi.org/10.1016/j.tate.2013.02.003

Upadyaya, K., Viljaranta, J., Lerkkanen, M.-K., Poikkeus, A.-M., & Nurmi, J.-E. (2012). Cross-lagged relations between kindergarten teachers' causal attributions, and children's interest value and performance in mathematics. *Social Psychology of Education*, *15*(2), 181–206. https://doi.org/10.1007/s11218-011-9171-1

Urdan, T. (2004). Predictors of academic self-handicapping and achievement: examining achievement goals, classroom goal structures, and culture. *Journal of Educational Psychology, 96*(2), 251–264. https://doi.org/10.1037/0022-0663.96.2.251

Urdan, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology, 44*(5), 331–349. https://doi.org/10.1016/j.jsp.2006.04.003

Urdan, T., & Turner, J. C. (2005). Competence motivation in the classroom. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 297–317). Guilford Publications.

Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study: Qualitative descriptive study. *Nursing & Health Sciences*, *15*(3), 398–405. https://doi.org/10.1111/nhs.12048

Van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & van Aken, M. A. G. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child Development*, *85*(3), 842–860. https://doi.org/10.1111/cdev.12169

Van den Noortgate, W., Opdenakker, M.-C., & Onghena, P. (2005). The effects of ignoring a level in multilevel analysis. *School Effectiveness and School Improvement*, *16*(3), 281–303. https://doi.org/10.1080/09243450500114850

Vansteenkiste, M., Sierens, E., Goossens, L., Soenens, B., Dochy, F., Mouratidis, A., Aelterman, N., Haerens, L., & Beyers, W. (2012). Identifying configurations of perceived teacher autonomy support and structure: Associations with self-regulated learning, motivation and problem behavior. *Learning and Instruction*, *22*(6), 431–439. https://doi.org/10.1016/j.learninstruc.2012.04.002

Vedder-Weiss, D., & Fortus, D. (2017). Teachers' mastery goals: Using a self-report survey to study the relations between teaching practices and students' motivation for science learning. *Research in Science Education, 48,* 181–206. https://doi.org/10.1007/s11165-016-9565-3

Voelkl, K. E. (1995). School warmth, student participation, and achievement. *Journal of Experimental Education, 63*(2), 127–138. https://doi.org/10.1080/00220973.1995.9943817

Vogt, D. S., King, D. W., & King, L. A. (2004). Focus groups in psychological assessment: Enhancing content validity by consulting members of the target population. *Psychological Assessment*, *16*(3), 231–243. https://doi.org/10.1037/1040-3590.16.3.231

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*(4), 548–573. https://doi.org/10.1037/0033-295X.92.4.548

Weiner, B., Frieze, I. H., Kukla, A., Reed, L., Rest, S., & Rosenbaum, R. M. (1971). *Perceiving the causes of success and failure.* General Learning Press.

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology, 89*(3), 451–469. https://doi.org/10.1037/0022-0663.89.3.451

Wiles, R. (2013). *What are qualitative research ethics*? Bloomsbury.

Wilkins, J. L. M. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education*, *11*(2), 139–164. https://doi.org/10.1007/s10857-007-9068-2

Wolff Smith, L. J., & Beretvas, S. N. (2014). The impact of using incorrect weights with the multiple membership random effects model. *Methodology*, *10*(1), 31–42. https://doi.org/10.1027/1614-2241/a000066

Wolters, C. (2004). Advancing achievement goal theory: Using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *Journal of Educational Psychology, 96*(2)*,* 236–250. https://doi.org/10.1037/0022-0663.96.2.236

Wolters, C. A., & Daugherty, S. G. (2007). Goal structures and teachers' sense of efficacy: Their relation and association to teaching experience and academic level. *Journal of Educational Psychology*, *99*(1), 181–193. https://doi.org/10.1037/0022-0663.99.1.181

Wolters, C. A., Fan, W., & Daugherty, S. (2010). Teacher-reported goal structures: Assessing factor structure and invariance. *The Journal of Experimental Education*, *79*(1), 1–29. https://doi.org/10.1080/00220970903292934

Yeager, D. S., & Dweck, C. S. (2020). What can be learned from growth mindset controversies? *American Psychologist, 75*(9), 1269–1284. https://doi.org/10.1037/amp0000794

Yerrick, R., Parke, H., & Nugent, J. (1997). Struggling to promote deeply rooted change: The "filtering effect" of teachers' beliefs on understanding transformational views of teaching science. *Science Education*, *81*(2), 137–159. https://doi.org/10.1002/(SICI)1098-237X(199704)81:2<137::AID-SCE2>3.0.CO;2-G

Yu, J., & McLellan, R. (2019). Beyond academic achievement goals: The importance of social achievement goals in explaining gender differences in self-handicapping. *Learning and Individual Differences*, *69*, 33–44. https://doi.org/10.1016/j.lindif.2018.11.010

Zander, L., Brouwer, J., Jansen, E., Crayen, C., & Hannover, B. (2018). Academic self-efficacy, growth mindsets, and university students' integration in academic and social support networks. *Learning and Individual Differences*, *62*, 98–107. https://doi.org/10.1016/j.lindif.2018.01.012

Zhou, J., & Urhahne, D. (2013). Teacher judgment, student motivation, and the mediating effect of attributions. *European Journal of Psychology Education, 28*, 275–295. https://doi.org/10.1007/s10212-012-0114-9

Zhu, M., Urhahne, D., & Rubie-Davies, C. M. (2018). The longitudinal effects of teacher judgement and different teacher treatment on students' academic outcomes. *Educational Psychology*, *38*, 648–668. https://doi.org/10.1080/01443410.2017.1412399

Appendix A

Questionnaires Administered in the Main Study

A.1 Teacher motivational beliefs questionnaire

Table A.1

Teacher Motivational Beliefs Questionnaire

Scale	Original item	Adapted item	Item ID
Fixed mindset	[a]You have a certain amount of intelligence, and you can't really do much to change it.	3. My pupils have a certain amount of ability at maths, and they can't really do much to change it.	TF1
	[a]Your intelligence is something about you that you can't change very much.	6. My pupils' ability at maths is something about them that they can't change very much.	TF2
	[a]To be honest, you can't really change how intelligent you are.	20. To be honest, my pupils can't really change how good at maths they are.	TF3
	[a]You can learn new things, but you can't really change your basic intelligence.	22. My pupils can learn new things, but they can't really change their basic ability at maths.	TF4
Growth mindset	[a]No matter who you are, you can significantly change your intelligence level.	12. No matter who my pupils are, they can significantly change their ability at maths.	TG1
	[a]You can always substantially change how intelligent you are.	8. My pupils can always substantially change how good at maths they are.	TG2
	[a]No matter how much intelligence you have, you can always change it quite a bit.	14. No matter how much ability at maths my pupils have, they can always change it quite a bit.	TG3
	[a]You can change even your basic intelligence level considerably.	17. My pupils can change even their basic ability at maths considerably.	TG4
Mastery goal	[b]It is important for me to continue to learn more about teaching.	19. It is important for me to continue to learn more about teaching maths.	TM1
	[b]One of my goals for teaching is to develop more effective teaching methods.	4. One of my goals for teaching maths is to develop more effective teaching methods.	TM2
	[b]The development of my teaching abilities is important to me.	13. The development of my teaching abilities in maths is important to me.	TM3

Performance-approach goal	[b]It is important to me to be praised for having higher teaching abilities than other teachers.	2. It is important to me to be praised for having higher teaching abilities in maths than other teachers.	TAp1
	[b]It is important that my students think that I am a better teacher than other teachers in the school.	10. It is important that my pupils think that I am a better maths teacher than other teachers in the school.	TAp2
	[b]One of my goals for teaching is to be recognized as one of the best teachers in the school.	15. One of my goals for teaching maths is to be recognised as one of the best maths teachers in the school.	TAp3
	[b]It is important for me to teach better than other teachers.	16. It is important for me to teach maths better than other teachers.	TAp4
Performance-avoidance goal	[b]It is important that my teaching abilities are not inferior to that of most of my colleagues.	7. It is important that my teaching abilities in maths are not inferior to that of most of my colleagues.	TAv1
	[b]I don't want to look like an incompetent teacher to my fellow teachers.	5. I don't want to look like an incompetent maths teacher to my fellow teachers.	TAv2
	[b]I don't want to show poor teaching skills when the principal or parents observe one of my lessons.	9. I don't want to show poor maths teaching skills when the head teacher or members of the leadership team observe one of my lessons.	TAv3
Success-to-ability attribution	[c]When the child does well in some assignment in kindergarten, it is probably because he / she has ability.	11. If a pupil does well in maths, it is probably because he / she has a high ability at maths.	TS1
Success-to-effort attribution	[c]When the child does well in some assignment in kindergarten, it is probably because he / she tries hard.	1. If a pupil does well in maths, it is probably because he / she tries hard.	TS2
Failure-to-ability attribution	[c]When the child does not do well in some kindergarten assignments, it is probably because he / she lacks ability.	18. If a pupil does poorly in maths, it is probably because he / she lacks ability at maths.	TF1
Failure-to-effort attribution	[c]When the child does not do well in some kindergarten assignments, it is probably because he / she has not invested effort.	21. If a pupil does poorly in maths, it is probably because he / she has not invested enough effort.	TF2
Perceived ability/ expectation	[d]If you were to list all the students in your class from the worst to the best in maths, where would you put yourself?	23. If you were to list all the pupils in the class from the lowest achieving to the highest achieving in maths, where would you put [Pupil]?	TEx1
	[e]How well do you expect this child to do next year in [domain]?	24. How well do you expect [Pupil] to perform this year in maths?	TEx2

Note. Responses to all items except for perceived ability/expectation were given on a 5-point Likert-type scale (0 = *Strongly disagree*, 1 = *Disagree*, 2 = *Neither disagree nor agree*, 3 = *Agree*, 4 = *Strongly agree*). Responses to item TEx1 were given on a 5-point Likert-type scale (0 = *One of the lowest achieving*, 1 = *Below average*, 2 = *Average*, 3 = *Above average*, 4 = *One of the highest achieving*). Responses to item TEx2 were given on a 5-point Likert-type scale (0 = *Not at all well*, 1 = *Not well*, 2 = *OK*, 3 = *Well*, 4 = *Very well*). Items were taken from: [a]Theories of Intelligence Scale by Dweck (1999); [b]Goals for Teaching Scale by Shim et al. (2013); [c]Upadyaya et al. (2012); [d]Eccles et al. (1993); [e]Wigfield et al. (1997).

A1.1 Demographic questions

D1. What is your gender? (Male, Female, Other)
D2. For how many years have you been working as a teacher? (open answer)
D3. For how many years have you taught maths (including the current year)? (open answer)
D4. What subject/s in addition to maths are you teaching? (open answer)
D5. What is the highest level of education that you have completed? (Bachelor’s Degree, PGCE, Master’s Degree, Doctoral Degree)

A.2 Student motivational beliefs questionnaire

Table A.2

Student Motivational Beliefs Questionnaire

Scale	Original item	Adapted item	Item ID
Fixed mindset	[a]I don't think I personally can do much to increase my intelligence	17. I don't think I can do much to increase my ability at maths.	PF1
	[a]My intelligence is something about me that I personally can't change very much.	15. My ability at maths is something about me that I can't change very much.	PF2
	[a]To be honest, I don't think I can really change how intelligent I am.	26. To be honest, I don't think I can really change how good at maths I am.	PF3
	[a]I can learn new things, but I don't have the ability to change my basic intelligence.	7. Even though I can learn new things in maths, I can't change my basic ability at maths.	PF4
Growth mindset	[a]With enough time and effort I think I could significantly improve my intelligence level.	20. With enough time and effort, I think I could improve my ability at maths a lot.	PG1
	[a]I believe I can always substantially improve on my intelligence.	24. I believe I can always improve my ability at maths a lot.	PG2
	[a]Regardless of my current intelligence level, I think I have the capacity to change it quite a bit.	3. No matter what my ability at maths is now, I think I can change it quite a bit.	PG3
	[a]I believe I have the ability to change my basic intelligence level considerable over time.	11. I believe I can change my basic ability at maths a lot over time.	PG4
Mastery goal	[b]It's important to me that I learn a lot of new concepts this year.	22. It's important to me that I learn a lot of new maths concepts this year.	PM1
	[b]One of my goals in class is to learn as much as I can.	2. One of my goals in my maths lessons is to learn as much as I can.	PM2
	[b]One of my goals is to master a lot of new skills this year.	14. One of my goals in maths is to master a lot of new skills this year.	PM3
	[b]It's important to me that I thoroughly understand my class work.	19. It's important to me that I completely understand my maths work.	PM4
	[b]It's important to me that I improve my skills this year.	9. It's important to me that I improve my maths skills this year.	PM5
Performance-approach goal	[b]It's important to me that other students in my class think I am good at my class work.	1. It's important to me that other pupils in my class think I am good at my maths work.	PAp1

	[b]One of my goals is to show others that I'm good at my class work.	6. One of my goals in maths is to show others that I'm good at my maths work.	PAp2
	[b]One of my goals is to show others that class work is easy for me.	12. One of my goals in maths is to show others that maths work is easy for me.	PAp3
	[b]One of my goals is to look smart in comparison to the other students in my class.	16. One of my goals in maths is to look smart in comparison to the other pupils in my class.	PAp4
	[b]It's important to me that I look smart compared to others in my class.	23. It's important to me that I look smart compared to others in my maths class.	PAp5
Performance-avoidance goal	[b]It's important to me that I don't look stupid in class.	4. It's important to me that I don't look stupid in my maths lessons.	PAv1
	[b]One of my goals is to keep others from thinking I'm not smart in class.	10. One of my goals in maths is to keep others from thinking I'm not smart at maths.	PAv2
	[b]It's important to me that my teacher doesn't think that I know less than others in class.	13. It's important to me that my maths teacher doesn't think that I know less than others.	PAv3
	[b]One of my goals in class is to avoid looking like I have trouble doing the work.	18. One of my goals in my maths lessons is to avoid looking like I have trouble doing the work.	PAv4
	-	27. One of my goals in maths is that I don't look stupid compared to others in my class. (later excluded)	PAv5
Success-to-ability attribution	[c]When the child does well in some assignment in kindergarten, it is probably because he / she has ability.	21. If I do well in maths, it's probably because I have a lot of ability at maths.	PS1
Success-to-effort attribution	[c]When the child does well in some assignment in kindergarten, it is probably because he / she tries hard.	8. If I do well in maths, it's probably because I try hard.	PS2
Failure-to-ability attribution	[c]When the child does not do well in some kindergarten assignments, it is probably because he / she lacks ability.	5. If I do poorly in maths, it's probably because I don't have a lot of ability at maths.	PF1
Failure-to-effort attribution	[c]When the child does not do well in some kindergarten assignments, it is probably because he / she has not invested effort.	25. If I do poorly in maths, it's probably because I don't try hard enough.	PF2
Self-concept	[d]Compared to other students, how well do you expect to do in math this year?	28. Compared to other pupils, how well do you expect to do in maths [this/next] year?	PSC1
	[d]How well do you think you will do in your math course this year?	29. How well do you think you will do in your maths class [this/next] year?	PSC2
	[d]How good at math are you?	30. How good at maths are you?	PSC3
	[d]If you were to order all the students in your math class from the worst to the best in maths, where would you put yourself?	31. If you would order all the pupils in your maths class from the worst to the best in maths, where would you put yourself?	PSC4

[d]How have you been doing in math this year?	32. How have you been doing in maths [last/this] year?	PSC5

Note. Responses to all items except for self-concept were given on a 5-point Likert-type scale (0 = *Not at all true*, 1 = *Not true*, 2 = *Neither true nor untrue*, 3 = *True*, 4 = *Very true*). Response scales for self-concept items were for PSC1 (0 = *Much worse*, 1 = *A little bit worse*, 2 = *Similar*, 3 = *A little bit better*, 4 = *Much better*), PSC2 (0 = N*ot at all well*, 1 = *Not so well*, 2 = *OK*, 3 = *Well*, 4 = *Very well*), PSC3 (0 = *Not at all good*, 1 = *Not good*, 2 = *Neither good nor bad*, 3 = *Good*, 4 = *Very good*), PSC4 (0 = *One of the worst*, 1 = *Worse than many*, 2 = *In the middle*, 3 = *Better than many*, 4 = *One of the best*), and PSC5 (0 = *Not at all well*, 1 = *Not so well*, 2 = *OK*, 3 = *Well*, 4 = *Very well*). The items were taken from: [a]Implicit Theories of Intelligence Self-Theory scale by De Castella and Byrne (2015); [b]Personal Achievement Goal Orientations (Revised) by Midgley et al. (2000); [c]Upadyaya et al. (2012); [d]Ability/Expectancy Subscale by Eccles et al. (2005).

A.2.1 Demographic questions

D1. My Name is: (open response)
D2. I am a boy / girl (circle which one)
D3. I am in maths class: ________(for example Set 1, 7a, Middle Set).
D4. My Maths teacher is: (open response)

A.3 Fixed-ability practices questionnaires for teachers and students

Table A.3

Fixed-Ability Practices Questionnaire for Teachers and Students

Scale	Original item	Teacher item	Pupil item	ID
Tasks				
Unsophisticated task differentiation	-	27. I give low achieving pupils easier tasks than others.	27. My teacher gives easier tasks to pupils who don't do well in maths.	TD1
	-	18. I give low achieving pupils less work than others.	18. My teacher gives less work to pupils who don't do well in maths.	TD2
	-	7. I set high achieving pupils more difficult work than others.	7. Pupils who do well in my maths class get harder tasks than others.	TD3
	-	33. I assign struggling pupils fewer tasks than others.	33. Pupils who struggle in maths get fewer tasks to work on.	TD4
Unidimensional tasks	-	4. I assign tasks that need to be solved in one specific way.	9. My teacher gives us questions that can be answered in many different ways. (r)	TU1
	-	37. I assign tasks that have only one correct solution.	37. My teacher gives us tasks that have only one right answer.	TU2
	-	52. I ask my pupils to explain how they get their answers. (r)	52. My teacher asks us to explain how we get our answers. (r)	TU3
	-	24. When I set work, pupils can decide whether they want to do the more challenging extension tasks. (r)	24. When my teacher sets us work, we can choose to do more challenging extension tasks. (r)	TU4
	[e]In this subject the tasks are designed to be rich in variety.	13. I provide a variety of different learning activities during lessons. (r)	13. My teacher gives us a mixture of different learning activities in our lessons. (r)	TU5
	-	34. I ask my pupils to complete tasks in the same order.	34. My teacher wants all pupils to complete tasks in the same order.	TU6
Grouping				
Ability grouping	[a]When I arrange the students into groups, I do it according to varied abilities in each group.	44. I put pupils of similar abilities together when I organise pair or group work.	44. For group or pair work, my teacher puts pupils with a similar ability at maths together.	GA1

	[a]When I arrange the students into groups, I do it according to varied abilities in each group.	51. I put pupils of varied abilities together when I organise pair or group work. (r)	51. My teacher wants pupils who do well in maths to work with pupils who don't do as well. (r)	GA2
	-	32. I ask pupils to work together who perform at a similar level in maths.	32. For group or pair work, my teacher puts pupils who work at a similar level in maths together.	GA3
	-	6. I use mixed-ability groupings when I do group work during lessons. (r)	6. My teacher wants us to work in mixed-ability groups (pupils who do well in maths work with pupils who struggle). (r)	GA4
Stability	-	26. Over the course of the school year, I put the same pupils together for group work.	26. My teacher usually puts the same pupils together for group work.	GS1
	-	11. The same pupils usually work together when we do group work.	11. The same pupils usually work together when we do group work.	GS2
	-	36. I vary grouping arrangements so that different pupils can work together throughout the year. (r)	36. My teacher changes the groups for group work, so that different pupils can work together. (r)	GS3
	[e]In this subject the teacher watches to ensure that we don't work with the same classmates all the time.	42. I ensure that pupils don't work with the same classmates all the time. (r)	42. My teacher makes sure we don't work with the same classmates all the time. (r)	GS4
Evaluation				
Performance over process recognition	-	2. I praise pupils for putting a lot of effort into their maths work, even if they don't do everything correctly.	2. My teacher praises us for trying hard in maths even if we don't do everything correctly.	EP1
	[b1]I make a special effort to recognize students' individual progress, even if they are below grade level./ [b3]My teacher recognizes us for trying hard.	10. I make a special effort to recognise pupils for working hard.	10. My teacher points out pupils who work very hard in class.	EP2
	-	28. I emphasise to pupils that giving their best is more important than getting good grades.	28. My teacher says that giving our best is more important than getting good grades.	EP3
	[b1]I make a special effort to recognize students' individual progress, even if they are below grade level./ [e]In this subject the	48. I make a special effort to recognise pupils' individual progress.	48. My teacher points out when someone has improved in maths.	EP4

	teacher points out when someone has improved.			
	[e]In this subject we get feedback concerning our learning progress.	23. I give pupils feedback about their learning progress, for example when they are improving or getting worse.	23. My teacher gives us feedback about our progress, for example when we improve or get worse.	EP5
Social comparison	[b1]I point out those students who do well as a model for the other students./ [b3]My teacher points out those students who get good grades as an example to all of us.	17. I point out those pupils who do well in maths as a model for others.	17. My teacher points out pupils who get good maths grades, as an example to all of us.	ES1
	[b3]My teacher lets us know which students get the highest scores on a test.	9. I tell my class which pupils get the highest scores on a maths test	4. My teacher tells us which pupils get the highest scores on a maths test.	ES2
	[b1]I help students understand how their performance compares to others./ [b2]My teacher tells us how we compare to other students.	1. I help pupils understand how their maths performance compares to others.	12. My teacher tells us how we compare to other pupils in maths (for example, how fast we are and how many correct answers we get).	ES3
	-	30. When I assess my pupils' maths work or discuss their performance, I do so in front of the class.	30. My teacher talks about how well pupils do in maths in front of the whole class.	ES4
Little autonomy				
Little choice	[d]When I am doing something that interests me - T gives me enough time to finish it.	22. When pupils do something that interests them, I ensure I give them enough time to finish it. (r)	22. When I do something that interests me, my teacher gives me enough time to finish it. (r)	AC1
	[d]T allows me to choose how to do my work in the classroom.	38. I allow my pupils to choose how they do their work in the classroom. (r)	38. My teacher allows me to choose how to do my work in the classroom. (r)	AC2
	[d]When T gives us an assignment she allows us to choose which questions to answer.	15. When I set work, my pupils can choose between several questions or tasks. (r)	15. When my teacher sets us work, we can choose which tasks to do. (r)	AC3
	[d]T encourages me to work in my own way.	45. I let my pupils work in their own way. (r)	45. My teacher lets me work in my own way. (r)	AC4
	[a]I allow my students to decide whether they work individually or in groups (in class and at home).	31. I allow my pupils to decide whether they work on their tasks individually or with others. (r)	31. My teacher allows us to decide whether we work alone or with others. (r)	AC5
Little promotion of self-regulation	[a]The evaluation of students' work is performed by the students themselves (self-evaluation).	43. I ask my pupils to assess their own maths work. (r)	43. My teacher asks us to check and correct our own work. (r)	ASR1

	-	8. I ask my pupils to think about the areas in maths that they need to improve. (r)	8. My teacher asks us to think about the areas in maths that we need to improve. (r)	ASR2
	[e]In this subject we should set learning goals for ourselves.	35. I ask my pupils to set their own learning goals. (r)	35. My teacher asks us to set our own learning goals. (r)	ASR3
	[e]In this subject we should monitor our learning.	20. I ask my pupils to evaluate how well they are learning. (r)	20. My teacher asks us to evaluate how well we are learning. (r)	ASR4
	-	49. I ask my pupils to evaluate their own progress in maths. (r)	49. My teacher asks us to evaluate our progress in maths. (r)	ASR5
Differential value	-	21. I carefully listen to the ideas and contributions of all my pupils. (r)	21. My teacher seems to listen more carefully to pupils who do well in maths.	DV1
	-	14. I tend to ask pupils who do well in maths to give answers in front of the class more often than others.	14. My teacher asks pupils who do well in maths to give answers in front of the class more often than others.	DV2
	-	3. I tend to spend more time with pupils who are good at maths than with others.	3. My teacher spends more time with pupils who are good at maths than with others.	DV3
	-	47. I tend to be more interested in pupils who are good at maths than in others.	47. My teacher shows more interest in pupils who are good at maths than in others.	DV4
	-	39. I tend to give less attention to pupils who don't do well in maths.	39. My teacher gives less attention to pupils who don't do well in class.	DV5
Little promotion of risk-taking				
Little encouragement of help-seeking	[a]I encourage my students to help each other and to get help from each other during class work.	5. I encourage my pupils to help each other and to get help from each other during maths lessons. (r)	5. My teacher tells us to help each other during lessons. (r)	RH1
	[c]This teacher answers questions carefully and thoroughly.	41. I answer questions from my pupils carefully and thoroughly. (r)	41. My teacher answers questions carefully and thoroughly. (r)	RH2
	[c]This teacher provides sufficient time for students to ask questions.	50. I provide a lot of time for pupils to ask questions. (r)	50. My teacher gives us a lot of time to ask questions. (r)	RH3
	[c]This teacher encourages students to ask for help, even after class.	19. I encourage pupils to ask for help any time, even after class. (r)	19. My teacher allows us to ask for help any time, even after class. (r)	RH4
Unconstructive handling of mistakes	-	40. When pupils make a mistake, I emphasise what they can learn from it. (r)	40. When someone makes a mistake, my teacher shows us how to learn from it. (r)	RM1
	-	25. When pupils make mistakes, I give them time to try again. (r)	25. When we make mistakes, my teacher gives us time to try again. (r)	RM2

-	12. I get annoyed or disappointed when pupils make mistakes.	1. My teacher looks annoyed or disappointed when we make mistakes.	RM3
-	46. I show pupils my disapproval when they make mistakes.	46. My teacher tells us off when we make a mistake.	RM4
-	16. When a pupil makes a mistake, I try not to give it too much attention but to quickly move on.	16. When we make a mistake, my teacher ignores it and quickly moves on.	RM5
-	29. When a pupil makes a mistake, I stop others from making fun of it. (r)	29. When we make a mistake, my teacher stops others from making fun of it. (r)	RM6

Note. R = reversed. Responses to all items were given on a 5-point Likert-type scale (0 = *Never*, 1 = *Almost never*, 2 = *Occasionally* (for teachers) /*Sometimes* (for students), 3 = *Often*, 4 = *Always*). The items were taken from: [a]TARGETS-Related Mastery-Emphasizing Practices by Vedder-Weiss and Fortus (2017); [b]PALS (1=Approaches to Instruction, 2 = Perception of Classroom Goal Structures, 3 = Perception of Teacher's Goals) by Midgley et al. (2000); [c]Perceived Teacher Support of Questioning Scale by Butler and Shibaz (2008); [d]Providing Choice by Assor et al. (2002); [e]TARGET Goal Structure by Lüftenegger et al. (2017).

Appendix B

Example Interview Schedules

B.1 Example teacher interview schedule

Introduction

- Thanks for all your help during the last weeks
- Purpose: to understand the reasons why you do or say certain things in a given situation
- Discuss different practices that I observed, sometimes show example video, and ask you what your reasons were for doing or saying a given thing (lots of "why" questions)
- Ensure that the teacher seems comfortable and at ease. Check if they are happy to discuss their reasons for engaging in different practices
- Permission to record?

Warm-up questions

Q1: How did you feel about the observations in general?
Q2: Do you think you acted differently? Do you think the pupils acted differently?

Main Part 1: Discussion of practices

Dimension	**Question**
Tasks **Unsophisticated task differentiation**	• Do you always give all pupils the same tasks? Why/Why not? • Does it happen that you give some pupils easier work than others? Why/Why not? • Would you determine who does what or let them choose? • Do you sometimes give less work to pupils who are not doing well? Why/Why not?
Tasks **Unidimensional tasks**	*You sometimes set questions for which pupils can use different ways of solving them. For instance, they had to shade a fraction of a grit, and then a fraction of that fraction to get 1/8. So there were different ways of getting 1/8.* • What do you think about these more open-ended tasks compared to tasks where there is one clear way of doing it and one clear answer that is correct? • Do you teach pupils different methods for solving the same maths problem? Why/Why not? • When you give students tasks, do you want that they all do them in the same order? Why/Why not? • When pupils give you their answer to a question, would you say you ask them to explain how they got their answer? Why/Why not? • You usually put a Star Question on the board. What is that and why is it being used?

	• Would you say that you challenge your pupils a lot? Why/Why not?
Grouping	• Can you tell me why pupils are sitting together at a table the way they are? Who made the decision and why? • Do you look at their ability when deciding who should work together? Why/Why not? • Do you ever change how they sit or is this stable throughout the year? Why? • In general, can students change with whom they want to work? Why/why not? • While I was here, I did not see any group work. Do you sometimes do group work or pair work? Why/Why not? • If you do group work, how do you determine who works together and why? • Do you usually put the same pupils together and why?
Evaluation **Performance over process recognition**	*There was one situation in which you pointed out that students had made big progress from one lesson to the next.* **<u>Show CLIP 3</u>** • Why did you point this out? • Would you say that you give your students regular feedback about their performance, if so why? • In general, would you say that you praise your students and if so, what would you say you usually praise them for and why? • Would you say that you recognise your students for their effort? Would you comment on it if a student worked particularly hard? • Do you take their effort into account when you give them grades or are these only based on actual performance? Why?/Why not? • Students can get merits or demerits. What do you give these usually for? • Do you reward them for things like getting many correct answers? Why/Why not? • When you give merits or demerits do you do this in front of the class? Why/Why not? *After the knowledge quiz on Friday, you gave a maths clinic slip to students who performed worse than the week before.* **<u>Show CLIP 10</u>** • You said you do not want to give demerits or correction but rather the chance to improve. Why did you do this? Do you sometimes give demerits for declining performance? Why/why not?
Evaluation	*There were a couple of situations in which you asked your students to*

Social comparison	*indicate how many correct answers they got on a task* **<u>Show CLIP 5</u>** **<u>Show CLIP 9</u>** • Why are you doing this? • Are you using this as well to create some kind of competition between students? *After the knowledge quiz on the Friday I observed, you asked the students to read out the score they got.* **<u>Show CLIP 8</u>** • Why did you do this? Why did you decide to ask this in front of the whole class rather than asking them privately? • Do you speak about pupils' performance in front the class, such as how well they are doing or what scores they got on a test? Why/Why not?
Autonomy **Choice**	*I have noticed that in your class, and this school more generally, things are very structured. Classes usually begin in the same way, everybody starts working at tasks at the same time and they have a set amount of time to work on questions etc. Overall, pupils follow your directions very closely.* • Do you think this kind of structure and following instructions is important and if so why? • Are there also aspects in your lessons that you give your pupils choice over and why? Where do you prefer to make the decisions? Why? • Would you say that pupils in your class can work in their own way, and if so, why?
Autonomy **Promotion of self-regulation**	• Would you say that you ask pupils to set their own learning goals? Why/Why not? • Would you say that you ask pupils to engage in self-evaluation? Why/Why not? • What is your opinion on students marking or checking their own maths work? Do you do this? Why/why not?
Differential value	*Sometimes it seems that you pick students with their hands up and sometimes you pick them although they did not raise their hand and even tell them you pick them randomly.* • What is the reason for this? *There was one situation in which you said you would not pick a student because she already answered many questions before.* **<u>Show CLIP 6</u>**

	• Why did you do this and why did you point it out to students? • Would you say that you call on all pupils equally or are there pupils you take more often? Why/ why not? • I know that this is a very difficult question, but would you say that you give equal attention to all your pupils?
Risk-taking **Help-seeking**	• Would you say that you encourage students to ask you questions if they do not understand something? Why/why not? • Are there situations in which you do not want that they ask questions or in which you get annoyed when they ask for help? *I saw one situation in which you asked [Student name] to explain a task to [Student name].* • In general, would you say you encourage students to work together? Why/why not?
Risk-taking **Mistakes**	*Now I'd like to speak more about what you do when a pupil makes a mistake.* • What would you say is your usual reaction and why? • What do you think if students make a mistake? *There are two situations I would like to show you in which students make a mistake and then I want to speak about why you reacted in these ways.* **Show CLIP 7** • How would you say you reacted to students' mistakes? So in this situation, you are shaking your head but you are not really discussing the mistake. Why? **Show CLIP 2** • Why did you think you should have just given the correct answer yourself?

Conclusion

- Check if the teacher still seems comfortable and OK with how the interview went
- Thank the teacher for participation

B.2 Example student group interview schedule

Introduction

- Welcome and thanks for participation
- Purpose: I want to understand what you think about different teaching practices that Miss [Teacher name] uses
- I will tell you about something she did in the lessons or show you a part of the video recordings and then I would like you to tell me what you think about these things and how they make you feel
- The rules: discuss freely, say whatever you want, you can disagree with each other, try not to interrupt each other, don't share what someone says with pupils outside the group
- Check if students are happy with all of this and seem comfortable to discuss these questions
- Permission to record?

Warm-up questions

Q1: How was it for you to be observed? How did you feel about it?
Q2: Did you notice any changes in Madam's behaviour? Anything Madam did differently from usual or would you say she was as she always is?

Main Part 1: Discussion of practices

Dimension	**Question**
Tasks **Unsophisticated task differentiation**	• Does Miss sometimes give some students easier and some students more difficult tasks? Why do you think she does this/why not? How do you feel about it? • If yes, does she decide what tasks you have to do? How do you feel about it? • How would you feel if you would always get easier tasks than others? What do you think Miss would think about you? • Does it sometimes happen that Miss gives less work to students who are not doing well? • How would you feel about that? Would it be fair?
Tasks **Unidimensional tasks**	• Would you say that you do a lot of different learning activities or are most of the tasks you get similar? Why do you think that is/is not? How do you feel about it? • Do you always have to do all your tasks in the same order? What do you think about that? *You sometimes got tasks in which you could all use different ways of solving them and still get it correct. For instance, you had to shade a fraction of a grit, and then a fraction of that fraction to get 1/8. There were different ways of getting 1/8.* • What do you think about these tasks compared to tasks where there is only one way of doing it and which have one correct answer? • Would you say that Miss wants you to explain how you get your

	answers or is she satisfied when you tell her your result? Why do you think she does this/why not? How do you feel about it? • How would you feel if Miss would only care about whether you got the answer right or wrong? • I have seen that there is sometimes a Star Question on the board. What is that for? How do you feel about those? • Would you say that Miss challenges you a lot? Why do you think that is/is not? How do you feel about it?
Grouping	• Can you tell me why you are sitting together at the tables the way you are? Who made that decision and why? • Do you think that Miss thinks about your maths performance when putting you together? (similar or different pupils working together)? • What would you think about similar pupils having to work together? Why would a teacher do that? • Why would they put different students together? What do you prefer and why? • Do you sometimes do group work? Why/Why not? What do you think about that? • Could you change where you sit if you wanted to? Does Miss change you around sometimes? • When you do group work, is it always the same pupils who are in one group? *I know that you can change classes every half term depending on how well you did in your ability group.* • What do you think about that? How would you feel if you would have to stay in the same group the whole year?
Evaluation **Performance over process recognition**	• What do you think is more important to Miss, that you get correct solutions and good grades or that you give your best? Why? How do you feel about that? • Do you think that Miss notices it when you try hard? Why/Why not? • Does Miss sometimes give you feedback about how well you are doing? Does she tell you when you improved or perform worse? How do you feel about that? *There was one situation in which Miss pointed out that you made a big improvement from one lesson to the next.* **<u>Show CLIP 3</u>** • What do you think about her telling you this? • When Miss gives you a merit or demerit, what are these usually for? • Does she reward you for being good at maths (e.g., getting many

	correct answers?) How would you feel about it? • Does Miss give you merits and demerits in front of the class? How do you feel about that? *After your knowledge quiz on Friday, Miss gave a maths clinic slip to students who performed worse than the week before.* **<u>Show CLIP 10</u>** • Miss said she does not give you a demerit or correction but rather gives you the chance to improve. What do you think about this? Does it sometimes happen that you get demerits for performing worse? How would you feel about that?
Evaluation **Social comparison**	*There were a couple of situations in which Miss asked you to indicate how many correct answers you got.* **<u>Show CLIP 5</u>** **<u>Show CLIP 9</u>** • How do you feel about her asking you this? • What do you think about her doing this in front of everyone else? Would you look around to see who else got them correct? • How do you think pupils feel who did not get it correct? • Is there competition in the class? *There was one situation after the knowledge quiz you did on Friday, where Miss asked you to read out the score you got.* **<u>Show CLIP 8</u>** • How do you feel about telling your score to everyone else? How do those feel who did not get a high score? • Does Miss ever tell you how good pupils are at maths in front of the whole class? For example, what scores you got on a test or who did the best or worse? How would you feel about Miss doing that?
Autonomy **Choice**	• Would you say that Miss gives you choice over how you are doing your work? Why? • What are some of the things you have choice over and what are things that you cannot decide? What do you think about that? • Does she let you work in your own way? What are examples? • Can you sometimes choose what questions to answer on a work sheet or do you always have to do the same as a class? How do you feel about it? • Miss sometimes lets you work in pairs and sometimes she wants that you do individual work. Why do you think she does this?

Autonomy **Promotion of self-regulation**	• Sometimes Miss asks you to mark and correct your own work. Why? Feelings? • Does Miss ask you to think about the areas in maths you need to improve or to evaluate your progress? Why would she do this and not just tell you herself? How do you feel about it? • Does Miss ask you to set your own learning goals? Why/why not? What do you think about it?
Differential value	• I've seen that sometimes, Miss picks pupils who have their hands up, and sometimes she picks pupils randomly. Why do you think she does this? What do you think about it? *In one situation she did not want to pick a student because she already shared a lot of her answers.* **<u>CLIP 6</u>** • Why do you think she does not pick her again? • Would you say that Miss picks all students equally to answer questions in front of the class or does she pick some students more often? If so, who and why? • How would you feel if Miss would pick those that are a bit surer of their maths work more often? • Do you think that Miss gives equal attention to all pupils? If not, who does she spend more time with and why? • How would you feel if she would spend more time with pupils who do well in maths?
Risk-taking **Help-seeking**	• Do you feel that you can ask Miss for help if you don't understand something? Why/why not? *In general, it seemed as though Miss asks you if you have questions after she explained something new.* • If you do have a question, would you say she spends quite a bit of time to explain things in detail? Why do you think she does this? How do you feel about that? • Does Miss ever get annoyed or angry when you ask for help? How would you feel if she would? • Does Miss want that you help each other? Why/Why not?
Risk-taking **Mistakes**	• When someone in class makes a mistake, how does Miss react? (e.g., does she explain mistakes, does she get angry, does she ask you to try again…) • How do you feel about these reactions? Why do you think she reacts in these ways? *I have seen that Miss usually shakes her head without saying anything when someone makes a mistake or that she picks another student.* **<u>Show CLIP 7</u>**

	• What do you think about that? Is it helpful? *There was another situation in which you struggled to find the right answer and thus made mistakes. After one student found the correct answer, Miss said she should have just told you the correct answer herself.* **<u>Show CLIP 2</u>** • Why does she think she should have told you the answer straight away? • In general, would you say that Miss helps you to learn when you make a mistake? E.g., Do you discuss why you made a mistake? • What does Miss think about mistakes (are they bad or can they help?)

Main Part 2: Perception of teacher mindset

Finally, I would like to ask a couple more general questions about what you think your teacher thinks about learning maths.

Some people think that how good you are at maths cannot be changed and others think you can improve your ability.

Q1: Do you think your teacher believes that you can change how good you are at maths? Why or why not? How does this make you feel?
Q2: Do you think your teacher has high expectations for you? Why/ Why not? How does this make you feel?

Conclusion

- Check if students still seem comfortable
- Thanks for participation
- Please don't tell others what you or others said here
- See you again for questionnaires

Appendix C
Pilot Studies

C.1 Pilot study 1

The first pilot study was conducted to test whether all proposed research methods and the selected or developed instruments could produce the data necessary to answer the research questions.

C.1.1 Participants and procedure

The student sample consisted of 60 girls from four Year 7 classes (aged 11-12 years) from an all-girls secondary school. The teacher sample consisted of 16 maths teachers recruited through word-of-mouth and six maths teachers recruited from the participating school (15 women, 7 men; aged <25 – 50–59). All participants provided written assent and students' parents were given the opportunity to opt students out of the study. Students completed the motivational beliefs questionnaires and FAPs questionnaires during lessons while teachers did so online. To pilot the qualitative methods, one class was selected by the school's head of maths to participate in the observation of two maths lessons. Video segments showing instances of FAPs were selected for the video-stimulated recall interviews. The teacher of that class was interviewed individually, and four volunteering students took part in a group interview. The interviews lasted approximately 30 minutes each.

C.1.2 Analysis, results, and discussion

C.1.2.1 Quantitative instruments

Only information relevant to the modification of scales is provided here. Ideally, a factor analysis should have been performed for FAPs because the factor structure has not yet been established. However, this was not possible due to the small sample size. A greater sample size in the main study ensured that the factor structure could be examined based on the student data before using inferential statistics. The reliability of each scale was examined using the internal consistency coefficient Cronbach's α. Reliability was considered acceptable if $\alpha > .70$ (Field, 2009). The statistics *alpha if item deleted* and *corrected item-total correlation* were examined to check whether the deletion of certain items would improve reliability. An item-total correlation of at least .3 is recommended (Field, 2009).

C.1.2.1.1 Motivational beliefs questionnaire

Cronbach's α, the range, mean, and standard deviation for the motivational beliefs scales are presented in Table C.1 for teachers and Table C.2 for students. For teachers, all scales but the performance-approach goal scale demonstrated good reliability. Deleting item TAp1[20] led to an acceptable reliability of α=.71. This was the only item assessing teachers' desire to get verbally acknowledged for their teaching skills rather than just aiming to appear better than others. Yet, as the scale has demonstrated good reliability in previous research (α=.81; Shim et al., 2013), it would have been premature to remove this item based on the reliability of this small sample. Accordingly, the teacher motivational beliefs questionnaire was not modified.

All student motivational belief scales demonstrated good internal reliability except for the performance-avoidance goal scale. No item could have been deleted to reach acceptable reliability. Items PAv1 and PAv2 had a problematic item-total correlation with .20 and .29, respectively. However, this scale has been widely used in previous research and has demonstrated acceptable or good reliability (e.g., Midgley et al., 2000). The present finding was thus difficult to explain, especially given that the only adjustments to the original items were made to address maths rather than general performance, which has been recommended by the authors (Midgley et al., 2000). The reliability may thus be sample-specific, which is why I decided to add a fifth item (*"One of my goals in maths is that I don't look stupid compared to others in my class"*) with the option to later exclude that item if sufficient reliability was reached using only original items. This was the only adaptation made to the student motivational beliefs questionnaire.

Table C.1

Descriptive Statistics and Internal Reliability for Teacher Motivational Beliefs Scales

Subscale (N items)	*α*	Range	*Mean (SD)*
Fixed mindset (4)	.81	0.00 – 1.75	0.87 (0.57)
Mastery goal (3)	.85	3.00 – 4.00	3.51 (0.45)
Performance-approach goal (4)	.67	0.75 – 3.50	2.00 (0.71)
Performance-avoidance goal (3)	.81	1.67 – 4.00	3.06 (0.73)
Perceived ability/expectation (2)	.89[a]	3.18 – 3.59	3.43 (0.18)[a]
Success-to-ability attribution (1)	NA	1.00 – 3.00	1.81 (0.91)
Success-to-effort attribution (1)	NA	1.00 – 4.00	2.68 (0.78)
Failure-to-ability attribution (1)	NA	0.00 – 2.00	1.09 (0.61)
Failure-to-effort attribution (1)	NA	1.00 – 4.00	2.36 (0.85)

Note. N = 22 for all scales except for fixed mindset where N = 21 and perceived ability/expectation where N = 4. Answer options ranged on a scale from 0 to 4.
[a]Based on four teachers who responded to two items per student.

[20] See Appendix A for the items.

Table C.2

Descriptive Statistics and Internal Reliability for Student Motivational Beliefs Scales

Subscale (N items)	α	Range	*Mean (SD)*
Fixed mindset (4)	.71	0.00 – 3.75	1.14 (0.72)
Mastery goal (5)	.83	1.20 – 4.00	3.36 (0.61)
Performance-approach goal (5)	.86	0.00 – 3.80	1.17 (0.91)
Performance-avoidance goal (4)	.51	0.25 – 3.25	1.67 (0.71)
Self-concept (5)	.89	0.60 – 3.80	2.53 (0.72)
Success-to-ability attribution (1)	NA	0.00 – 4.00	2.17 (1.02)
Success-to-effort attribution (1)	NA	1.00 – 4.00	3.05 (0.73)
Failure-to-ability attribution (1)	NA	0.00 – 4.00	1.69 (1.34)
Failure-to-effort attribution (1)	NA	0.00 – 4.00	2.29 (1.30)

Note. Valid N = 59 for all scales except for performance-approach goal and failure-to-ability attribution where N = 58. Answer options ranged on a scale from 0 to 4.

C.1.2.1.2 Fixed-ability practices questionnaire

Table C.3 and Table C.4 present the internal reliability, range, mean, and standard deviations for the FAPs scales for teachers and students, respectively. The reliabilities showed that both the teacher and student FAPs questionnaires needed major revisions. Modifications were based on internal reliability, alpha if item deleted, corrected item-total correlation, and theoretical considerations. All modifications and the reasons for these adaptations can be found in Table C.5.

Table C.3

Descriptive Statistics and Internal Reliability for Teacher Fixed-Ability Practices Scales

Scale (N items)	α	Range	Mean (SD)
Tasks			
Unsophisticated task differentiation (5)	.61	0.20 – 2.20	1.01 (0.50)
Unidimensional tasks (4)	.66	1.75 – 3.75	2.50 (0.57)
Grouping			
Ability grouping (3)	.15	0.33 – 2.33	1.81 (0.52)
Stability (2)	NA[a]	0.50 – 2.00	1.34 (0.51)
Evaluation			
Performance over process recognition (5)	.55	0.80 – 2.60	1.36 (0.53)
Social comparison (6)	.82	0.67 – 1.75	1.75 (0.71)
Little autonomy			
Little choice (6)	.35	1.33 – 3.00	2.21 (0.47)
Little promotion of self-regulation (2)	.80	0.50 – 4.00	1.50 (0.82)
Differential value (5)	NA[a]	0.80 – 1.80	1.13 (0.30)
Little promotion of risk-taking			
Little encouragement of help-seeking (3)	.66	0.00 – 1.00	0.56 (0.38)
Unconstructive handling of mistakes (2)	.47	0.00 – 2.00	0.91 (0.58)

Note. N = 16. Answer options ranged on a scale from 0 to 4. Higher scores indicate more frequent fixed-ability practices. [a]Due to negative average covariance.

Table C.4

Descriptive Statistics and Internal Reliability for Student Fixed-Ability Practices Scales

Scale (N items)	Valid n	α	Range	*Mean (SD)*
Tasks				
Unsophisticated task differentiation (5)	58	.71	0.00 – 2.00	0.58 (0.44)
Unidimensional tasks (4)	56	.31	1.25 – 3.50	2.39 (0.52)
Grouping				
Ability grouping (3)	56	.15	0.33 - 3.33	1.89 (0.69)
Stability (1)	57	[a]NA	0.00 – 4.00	1.09 (1.21)
Evaluation				
Performance over process recognition (5)	57	.32	0.20 – 3.20	1.71 (0.62)
Social comparison (6)	56	.62	0.00 – 3.33	1.17 (0.68)
Little autonomy				
Little choice (6)	58	.58	1.00 – 3.83	2.56 (0.65)
Little promotion of self-regulation (2)	56	.36	0.50 – 4.00	2.26 (0.73)
Differential value (5)	53	.50	0.00 – 3.20	1.41 (0.70)
Little promotion of risk-taking				
Little encouragement of help-seeking (4)	58	.76	0.00 – 3.50	1.50 (0.82)
Unconstructive handling of mistakes (2)	54	.33	0.00 – 3.50	1.70 (0.81)

Note. Answer options ranged on a scale from 0 to 4. Higher scores indicate more frequent fixed-ability practices. [a]Due to negative average covariance.

C.1.2.2 Qualitative instruments

C.1.2.2.1 Observations

Observations were piloted to determine whether this method was appropriate to capture the teaching practices of interest and to reveal any practical issues that needed to be resolved prior to the main study. This section focuses on the lessons learned from pilot observations. Starting with practical issues, it became apparent that I needed to arrange with schools that I could arrive in the classroom at least 10 minutes before the lessons start to set up the recording equipment. I also realised that it would be beneficial to use a smaller camera and simpler audio recording because the elaborative set up of a large camera attracted much attention from students. I will now address whether the practices of interest can be captured through observations.

Table C.5

Overview of the Modifications of Student Fixed-Ability Practices Items After Pilot Study 1

Scale	Pilot 1 items	Modifications	Adapted/new items
Tasks			
Unsophisticated task differentiation	In our maths lessons, all pupils have to work on the same topic. (r)	Delete as this is likely always the case.	-
	In our maths lessons, all pupils get the same task (e.g., the same work sheet). (r)	"Tasks" instead of "task."	In our maths lessons, all pupils get the same tasks (e.g., the same work sheet). (r)
	My maths teacher gives all pupils challenging tasks even if we struggle. (r)	Delete as the item does not relate clearly to the conceptualisation.	-
	My maths teacher gives pupils who are not doing so well easier tasks than others.	Simplify wording.	My teacher gives pupils who don't do well in maths easier tasks.
	My maths teacher gives pupils who are not doing so well less work than others.	Simplify wording.	My teacher gives pupils who don't do well in class less work.
	-	Add two new items to form complete scale.	Pupils who do well in my maths class get harder tasks than others.
	-		My teacher gives more difficult maths work to pupils who do well in class.
Unidimensional tasks	My maths teacher gives us tasks that need to be solved in one specific way.	Simplify wording.	My teacher gives us maths problems that need to be solved in only one way.
	My maths teacher gives us tasks that have one correct solution.	Simplify wording and add "only" to be more explicit.	My teacher gives us tasks that have only one right answer.
	My maths teacher tells us in what order we should complete our tasks.	Delete because all inter-item correlations are negative.	-
	My maths teacher tells us in how much time we should complete our tasks.	Delete as the item does not relate clearly to the conceptualisation.	-
	-	Add two new items to form complete scale.	My teacher wants us to try different ways to solve a maths problem. (reversed)
	-		My teacher gives us questions that can be answered in many different ways. (reversed)
Grouping			
Ability grouping	When we do group work, my maths teacher puts pupils with similar ability at maths into one group.	Simplify wording.	My teacher puts pupils with the same ability in maths together when we work in groups.

	My maths teacher lets us choose with whom we want to work. (r)	Delete as the item does not relate clearly to the conceptualisation.	-
	My maths teacher gives us group work or pair work. (r)	Delete as the item does not relate clearly to the conceptualisation.	-
	-	Add three new items to form complete scale.	For group work, my teacher pairs pupils who are good at maths with pupils who are not as good at maths. (r)
	-		For group work, my teacher puts pupils together who work at a similar level in maths.
	-		My teacher puts pupils who are good at maths together for group work.
Stability	During the school year, my maths teacher puts the same pupils together in one group.	Rephrase.	My teacher usually puts the same pupils together for group work.
	-	Add three new items to form complete scale.	The same pupils usually work together when we do group work.
	-		My teacher changes the groups for group work so that different pupils work together. (r)
	-		My teacher usually changes the pupils who work together during group activities. (r)
Evaluation			
Performance over process recognition	My maths teacher recognises us for trying hard. (r)	"Praises" instead of "recognises."	My teacher praises us for trying hard in maths. (r)
	My maths teacher gives us time to really explore and understand new ideas. (r)	Delete as the item does not relate clearly to the conceptualisation.	-
	When my maths teacher gives us back our work, we can hand it in again after we corrected or improved it. (r)	Delete as the item does not relate clearly to the conceptualisation.	-
	My maths teacher praises pupils for being good at maths.	"Clever" instead of "being good" as the latter is more ambiguous.	My teacher praises pupils for being clever at maths.
	My maths teacher praises pupils for working hard. (r)	-	My teacher praises pupils for working hard. (r)
	-	Add new item to form complete scale.	My teacher says that giving our best is more important than getting good grades. (r)
Social comparison	My maths teacher gives special rewards to pupils who do the best work.	-	My teacher gives special rewards to pupils who do the best work.

	My maths teacher points out those pupils who get good maths grades as an example to all of us.	Simplify wording.	My teacher points out pupils who get good maths grades, as an example to all of us.
	My maths teacher lets us know which pupils get the highest scores on a maths test.	Simplify wording.	My teacher tells us which pupils get the highest scores on a maths test.
	My maths teacher tells us how we compare to other pupils.	Add "in maths."	My teacher tells us how we compare to other pupils in maths.
	My maths teacher says that we pupils should compete with each other in maths.	Simplify wording.	My teacher says that we should compete with each other in maths.
	When my maths teacher assesses our work or discusses our performance, he/she usually does so in front of the class.	Simplify wording.	My teacher talks about how well pupils do in maths in front of the whole class.
Little autonomy			
Little choice	All pupils can choose between different questions when our maths teacher gives us tasks (e.g., tests and worksheets). (r)	The following items were replaced by items from the choice scale by Assor et al. (2002).	
	My maths teacher allows all pupils to choose how we want to present our results (e.g., written answers, drawings and tables). (r)		
	My maths teacher gives all pupils choice over how our work is assessed (e.g., using marks or verbal evaluation, who assesses the work and which criteria will be used). (r)		
	When we pupils are working in groups or in pairs, my maths teacher decides who will do what.		
	My maths teacher allows all of us pupils to decide where we will sit during lessons. (r)		
	My maths teacher allows all of us pupils to decide whether we work alone or in groups. (r)	Retain but simplify wording.	My teacher allows us to decide whether we work alone or with others. (r)
	-		When I do something that interests me, my teacher gives me enough time to finish it. (r)
	-		My teacher allows me to choose how to do my work in the classroom. (r)
	-		My teacher asks us if there are things we would like to change about our lessons. (r)

	-		When my teacher gives us tasks, we can choose which questions to answer. (r)
	-		My teacher lets me work in my own way. (r)
Little promotion of self-regulation	My maths teacher asks us to assess our own maths work. (r)	Adapt wording because "assess" may be understood too formally.	My teacher asks us to check and correct our own work. (r)
	My maths teacher asks us to assess each other's math work. (r)	Adapt wording because "assess" may be understood too formally.	My teacher asks us to check each other's maths work. (r)
	-	Add new item from Assor et al. (2002) to form complete scale.	My teacher gives us a say in how our maths work is assessed, for example, how we decide what is good work and what is not. (r)
	-	Add new item to form complete scale.	My teacher asks us to mark each other's answers during lessons.
Differential value			
	My maths teacher carefully listens to what we have to say, no matter how good at maths we are. (r)	Rephrase to address differentiation more directly.	My teacher seems to listen more carefully to pupils who do well in maths.
	My maths teacher asks all pupils to give answers in front of the class, not just those that belong to the best of the class. (r)	Rephrase to address differentiation more directly.	My teacher asks pupils who do well at maths to give answers in front of the class more often than others.
	My maths teacher tells us that we should only give helpful feedback and never say something mean about the work of others. (r)	Rephrase to simplify.	My teacher tells us off if we say something unkind about the work of others. (r)
	My maths teacher gives us extra homework or tests when we are misbehaving.	Delete as the item does not relate clearly to the conceptualisation.	-
	My maths teacher calls more often on pupils that do not pay attention or are misbehaving.	Delete as the item does not relate clearly to the conceptualisation.	-
	-	Add two new items to form complete scale.	My teacher takes our answers seriously no matter how well we do in maths. (r)
	-		When we make mistakes, my teacher stops others from making fun of it. (r)
Little promotion of risk-taking			
Little encouragement of help-seeking	My maths teacher tells us to help each other and to get help from each other during lessons. (r)	Simplify wording.	My teacher tells us to help each other during lessons. (r)
	My maths teacher answers questions carefully and thoroughly.	-	My teacher answers questions carefully and thoroughly. (r)

	My maths teacher gives us a lot of time to ask questions. (r)	-	My teacher gives us a lot of time to ask questions. (r)
	My maths teacher allows us to ask for help any time, even after class. (r)	-	My teacher allows us to ask for help any time, even after class. (r)
Unconstructive handling of mistakes	My maths teacher tells us that mistakes are okay as long as we are learning from them. (r)	Phrase as a more direct action.	When someone makes a mistake, my teacher shows us how to learn from it. (r)
	My maths teacher tells us that getting right solutions is very important.	Delete as this does not refer to handling of mistakes.	-
	My maths teacher gives us challenging tasks even if we make mistakes. (r)	Not administered in Pilot Study 1 due to mistake in questionnaire but the item does not directly reflect the conceptualisation.	-
	-	Add new items based on the conceptualisation of adaptive and maladaptive teacher responses to mistakes by Tulis (2013).	When we make mistakes, our teacher gives us time to try again. (r)
	-		My teacher discusses our mistakes with the whole class, so that everybody can learn from them. (r)
	-		My teacher looks annoyed or disappointed when we make mistakes.
	-		My teacher tells us off when we make a mistake.
	-		When we make a mistake, my teacher tells us the right answer straight away.
	-		When we make mistakes, my teacher stops others from making fun of it. (r)
	-		When we make a mistake in class, my teacher picks another pupil to correct it.
	-		When we make mistakes, my teacher ignores them and quickly moves on.

Note. R = reversed. The new teacher fixed-ability practices questionnaire items were modelled after the new student items.

Tasks. For the task dimension, most aspects could be observed well, especially whether all students got the same topic, tasks, and materials. However, it could not be observed why some students may have gotten different tasks. The same was true for the expected product and whether different approaches to tasks were allowed. For instance, it was very apparent that the teacher did not offer students different tasks but assigned all students one specific task at any given time, which he expected to be solved in one specific way. However, I observed one instance in which a student said she wanted to approach a task in a different way, and the teacher let her do so, even though it was clear that he did not appreciate it.

Grouping. Direct grouping practices were observed only during the second lesson. This shows that some teachers may not engage in any grouping practices during observations, which then needs to be discussed in interviews. It could not be observed why certain students were seated or grouped together and whether grouping structures changed over time. Again, these aspects thus needed to be addressed during interviews.

Evaluation. Most aspects of this dimension could be observed well, such as teacher comments about student effort, understanding, and performance. While many instances of praise occurred, it was quite difficult to ascertain whether it was process-, product- or person-oriented praise as the comments were usually rather general (e.g., "well done"). Almost no instance was observed where the teacher explicitly praised effort or the process. Only very few instances of criticism occurred but these were more specific and usually process-oriented (e.g., not listening, not trying). I realised that it might be difficult to observe whether the teacher tends to evaluate students in public or private because these instances occurred infrequently. Hence, this needed to be addressed in interviews.

Little autonomy. Practices relating to student autonomy could be observed quite well. For example, students had choice over how to present the work they produced. For instance, when presenting group work, students decided who presented and with what means. It was also apparent that students could decide whether to work alone or in pairs. Instances of self- or peer-assessment were observed frequently but self-evaluations might be more difficult to observe as they presumably appear less frequently.

Differential value. I quickly realised that it was very difficult to observe whether the teacher showed equal interest in or gave equal attention to all students, simply because of a lack of familiarity with the students and the class in general. Similarly, it was difficult to judge teachers' affect towards different students and even more so to detect whether or not this was based on students' performance. These aspects thus needed to be addressed during interviews.

It was also not possible to observe whether the teacher called on all students equally due to unfamiliarity with students and whether the teacher valued and considered contributions from all students equally. Nevertheless, it could be observed whether the teacher at least appeared to encourage multiple different students to participate, especially those who seemed distracted or did not volunteer answers themselves.

Little promotion of risk-taking. As students frequently made mistakes during lessons, the teacher's reactions to these could be observed well, such as whether he discussed mistakes or simply moved on to the next student. Practices relating to help-seeking were also easily observed. For instance, the teacher verbally encouraged students to ask questions and took much time to answer. More difficult to observe was whether students sought help from each other because individual student conversations could not be heard in the video-recordings. This aspect thus needed to be addressed during interviews.

C.1.2.2.2 Interviews

Before conducting interviews, I selected video segments showing the teaching practises of interest. I used the conceptualisation of FAPs (see Table 2.1, Chapter 2) to identify these instances in the recordings. While I initially planned to only include practices that clearly linked to the dimensions, I realised that the interviews were a good opportunity to discuss situations that were ambiguous and thus difficult to interpret. In addition, I noticed that not all practices needed video-stimulations in order to be discussed. For example, it was easier to describe than show that all students received the same task. Since not all practices of interest occurred during lessons and because some practices could not be captured on video, such as the bases for grouping, some practices needed to be discussed without video-stimulation.

The interview schedule needed modifications according to practices that could or could not be recorded during observations. It was thus not possible in the main study to use the same interview schedule for every class, but the general protocol seemed to be appropriate. It was only slightly adapted after piloting as it became apparent that certain questions did not produce valuable data. For instance, I planned to ask participants to describe every video segment but that felt unnatural and cost time without yielding interesting insights. I also realised that instead of only discussing teaching practices that teachers actually used, it was interesting to address how students would feel about the opposite practices. For example, the observed teacher did not discuss performance publicly, so I asked students what they would think about public evaluations and students readily referred to their experiences with other teachers.

C.1.3 Conclusions

Overall, the first pilot study demonstrated that the chosen methods could produce the data necessary to answer the research questions. The motivational beliefs questionnaires were only modified for students by adding a fifth item to assess performance-avoidance goals. However, it became apparent that substantial modifications needed to be made for the FAPs questionnaires. Accordingly, I decided to pilot the FAPs questionnaire a second time. This was done only with students due to sample size considerations. Only minor adaptations were necessary for the qualitative methods.

C.2 Pilot study 2

A second pilot study was conducted to further adapt the fixed-ability practices questionnaires.

C.2.1 Participants and procedure

The second pilot study was conducted at a co-educational secondary school in the East of England, which also participated in the main study. 116 students (boys = 51, girls = 55, not reported = 10; aged 11-12 years) from five Year 7 classes participated. Notably, these classes were taught by teachers who did not participate in the main study. All students provided written assent and their parents were given the opportunity to opt them out of the study. Pen-and-paper questionnaires were completed during a regular school lesson.

C.2.2 Analysis, results, and discussion

Table C.6 presents the internal reliability, range, mean, and standard deviation of the FAPs scales for students. Again, no factor structure was examined due to the insufficient sample size. Instead, internal reliability, alpha if item deleted, corrected item-total correlation, and theoretical considerations were used to further adapt the scales. The alpha statistics already indicate that some scales needed substantial revisions. All modifications made to the student FAPs items as well as the reasons for these adaptations can be found in Table C.7. The FAPs questionnaires were modified based on student data alone and the teacher scales were later modelled after the student scales.

Table C.6

Descriptive Statistics and Internal Reliability for Student Fixed-Ability Practices Scales in Pilot Study 2

Scale (N items)	Valid n	α	Range	*Mean (SD)*
Tasks				
Unsophisticated task differentiation (5)	105	.72	0.00 – 3.00	0.82 (0.63)
Unidimensional tasks (4)	107	.43	0.50 – 3.50	1.93 (0.55)
Grouping				
Ability grouping (4)	101	.21	0.00 – 3.00	1.40 (0.61)
Stability (4)	107	.60	0.25 – 4.00	1.91 (0.75)
Evaluation				
Performance over process recognition (4)	107	.23	0.00 – 3.00	1.25 (0.64)
Social comparison (6)	109	.56	0.00 – 2.00	0.65 (0.53)
Little autonomy				
Little choice (6)	104	.70	0.50 – 4.00	2.33 (0.66)
Little promotion of self-regulation (4)	100	.26	0.25 – 3.75	1.81 (0.58)
Differential value (4)	109	.37	0.00 – 2.75	0.88 (0.62)
Little promotion of risk-taking				
Little encouragement of help-seeking (4)	105	.76	0.00 – 4.00	1.14 (0.81)
Unconstructive handling of mistakes (9)	107	.68	0.22 – 3.00	0.91 (0.52)

Note. Answer options ranged on a scale from 0 to 4. Higher scores indicate more frequent fixed-ability practices.

C.2.3 Conclusions

As shown in Table C.7, the FAPs scales were substantially modified following the second pilot study. Some items were deleted because of poor inter-item correlations or poor theoretical fit, and new items were developed to reach an adequate number of items, to increase reliability, and ensure that enough items could be retained in the main study. Other items were merely rephrased to simplify the wording. A larger sample size in the main study allowed for the examination of the factor structure of the student questionnaire.

Table C.7

Overview of the Modifications of Student Fixed-Ability Practices Items After Pilot Study 2

Scale	Pilot 2 items	Modifications	Adapted/new items
Tasks			
Unsophisticated task differentiation	In our maths lessons, all pupils get the same tasks (e.g., the same work sheet). (r)	Delete. This practice is not necessarily adaptive because students could also work on different tasks based on their own choice.	-
	My teacher gives pupils who don't do well in maths easier tasks.	Simplify wording.	My teacher gives easier tasks to pupils who don't do well in maths.
	My teacher gives pupils who don't do well in class less work.	Simplify wording.	My teacher gives less work to pupils who don't do well in maths.
	Pupils who do well in my maths class get harder tasks than others.	-	Pupils who do well in my maths class get harder tasks than others.
	My teacher gives more difficult maths work to pupils who do well in class.	Delete as too similar to previous item.	-
		New item to balance scale.	Pupils who struggle in maths get fewer tasks to work on.
Unidimensional tasks	My teacher gives us maths problems that need to be solved in only one way.	Delete as this item, the third, and the fourth all assess whether students can solve tasks in different ways. I only retained the fourth item as it had higher correlations with the second item.	-
	My teacher gives us tasks that have only one right answer.	-	My teacher gives us tasks that have only one right answer.
	My teacher wants us to try different ways to solve a maths problem. (reversed)	Delete as this item, the first, and the fourth all assess whether students can solve tasks in different ways. I only retained the fourth item as it had higher correlations with the second item.	-
	My teacher gives us questions that can be answered in many different ways. (r)	-	My teacher gives us questions that can be answered in many different ways. (r)
	-	New item based on conceptualisation of multidimensional tasks by Boaler and Staples (2008).	My teacher asks us to explain how we get our answers. (r)

	-	New item to assess provision of extension tasks based on Boaler (2016).	When my teacher sets us work, we can choose to do more challenging extension tasks.
	-	New item based on Marshall and Weinstein (1984) and Boaler and Staples (2008).	My teacher gives us a mixture of different learning activities in our lessons. (r)
	-	New item based on Marshall & Weinstein (1984).	My teacher wants all pupils to complete tasks in the same order.
Grouping			
Ability grouping	My teacher puts pupils with the same ability in maths together when we work in groups.	Simplify wording, replace “same ability” by “similar ability” to be more inclusive, and also include partner work.	For group or pair work, my teacher puts pupils with a similar ability at maths together.
	For group work, my teacher pairs pupils who are good at maths with pupils who are not as good at maths. (r)	Completely rephrase because of negative correlation with all other items.	My teacher wants pupils who do well in maths to work with pupils who don’t do as well. (r)
	For group work, my teacher puts pupils together who work at a similar level in maths.	Include pair work.	For group or pair work, my teacher puts pupils who work at a similar level in maths together.
	My teacher puts pupils who are good at maths together for group work.	Include pair work.	My teacher puts pupils who are good at maths together for group or pair work.
	-	New item to increase reliability.	My teacher wants us to work in mixed-ability groups (pupils who do well in maths work with pupils who struggle).
Stability	My teacher usually puts the same pupils together for group work.	-	My teacher usually puts the same pupils together for group work.
	The same pupils usually work together when we do group work.	-	The same pupils usually work together when we do group work.
	My teacher changes the groups for group work so that different pupils work together. (r)	Add “can.”	My teacher changes the groups for group work, so that different pupils can work together. (r)
	My teacher usually changes the pupils who work together during group activities. (r)	Delete because of low inter-item correlations. Add a new item instead.	-
		New item to balance scale based on Lüftenegger et al. (2017).	My teacher makes sure we don't work with the same classmates all the time. (r)

Evaluation			
Performance over process recognition	My teacher praises us for trying hard in maths. (r)	Add a phrase to emphasise the contrast to performance recognition.	My teacher praises us for trying hard in maths even if we don't do everything correctly. (r)
	My teacher praises pupils for being clever at maths.	Delete because all inter-item correlations are negative.	-
	My teacher praises pupils for working hard. (r)	Adapt wording to differentiate the item more from the first item of the subscale.	My teacher points out pupils who work very hard in class. (r)
	My teacher says that giving our best is more important than getting good grades. (r)	-	My teacher says that giving our best is more important than getting good grades. (r)
	-	New item based on Lüftenegger et al. (2017).	My teacher points out when someone has improved in maths. (r)
	-	New item based on Lüftenegger et al. (2017).	My teacher gives us feedback about our progress, for example when we improve or get worse.
Social comparison	My teacher gives special rewards to pupils who do the best work.	Delete because of low inter-item correlations – improves reliability.	-
	My teacher points out pupils who get good maths grades, as an example to all of us.	-	My teacher points out pupils who get good maths grades, as an example to all of us.
	My teacher tells us which pupils get the highest scores on a maths test.	-	My teacher tells us which pupils get the highest scores on a maths test.
	My teacher tells us how we compare to other pupils in maths.	Add examples to explain what kind of comparison is asked for.	My teacher tells us how we compare to other pupils in maths (for example, how fast we are and how many correct answers we get).
	My teacher says that we should compete with each other in maths.	Delete because of low inter-item correlations – improves reliability.	-
	My teacher talks about how well pupils do in maths in front of the whole class.	-	My teacher talks about how well pupils do in maths in front of the whole class.

Little autonomy			
Little choice	My teacher allows us to decide whether we work alone or with others. (r)	-	My teacher allows us to decide whether we work alone or with others. (r)
	When I do something that interests me, my teacher gives me enough time to finish it. (r)	-	When I do something that interests me, my teacher gives me enough time to finish it. (r)
	My teacher allows me to choose how to do my work in the classroom. (r)	-	My teacher allows me to choose how to do my work in the classroom. (r)
	My teacher asks us if there are things we would like to change about our lessons. (r)	Delete because of low inter-item correlations. In addition, the item is very broad, and the practice probably occurs very infrequently.	-
	When my teacher gives us tasks, we can choose which questions to answer. (r)	Slightly rephrase for clarity.	When my teacher sets us work, we can choose which tasks to do. (r)
	My teacher lets me work in my own way. (r)	-	My teacher lets me work in my own way. (r)
Little promotion of self-regulation	My teacher asks us to check and correct our own work. (r)	-	My teacher asks us to check and correct our own work. (reversed)
	My teacher asks us to check each other's maths work. (r)	Delete and instead focus more on evaluation rather than assessment/marking.	-
	My teacher gives us a say in how our maths work is assessed, for example, how we decide what is good work and what is not. (r)	Delete because this seems rather difficult in maths.	-
	My teacher asks us to mark each other's answers during lessons.	Delete and instead focus more on evaluation rather than assessment/marking.	-
	-	New item to balance scale.	My teacher asks us to think about the areas in maths that we need to improve.
	-	New item based on Lüftenegger et al. (2017).	My teacher asks us to set our own learning goals.
	-	New item based on Lüftenegger et al. (2017).	My teacher asks us to evaluate how well we are learning.
	-	New item to balance scale.	My teacher asks us to evaluate our progress in maths.

Differential value			
	My teacher seems to listen more carefully to pupils who do well in maths.	-	My teacher seems to listen more carefully to pupils who do well in maths.
	My teacher asks pupils who do well at maths to give answers in front of the class more often than others.	-	My teacher asks pupils who do well in maths to give answers in front of the class more often than others.
	My teacher tells us off if we say something unkind about the work of others. (r)	Delete as the item does not clearly reflect the conceptualisation.	-
	My teacher takes our answers seriously no matter how well we do in maths. (r)	Delete as it is not clear what it would mean when a teacher does not take an answer seriously.	-
	-	Add new item to balance scale and increase reliability.	My teacher spends more time with pupils who are good at maths than with others.
	-	Add new item to balance scale and increase reliability.	My teacher shows more interest in pupils who are good at maths than in others.
	-	Add new item to balance scale and increase reliability.	My teacher gives less attention to pupils who don't do well in class.
Little promotion of risk-taking			
Little encouragement of help-seeking	My teacher tells us to help each other during lessons. (r)	-	My teacher tells us to help each other during lessons. (r)
	My teacher answers questions carefully and thoroughly. (r)	-	My teacher answers questions carefully and thoroughly. (r)
	My teacher gives us a lot of time to ask questions. (r)	-	My teacher gives us a lot of time to ask questions. (r)
	My teacher allows us to ask for help any time, even after class. (r)	-	My teacher allows us to ask for help any time, even after class. (r)
Unconstructive handling of mistakes	When someone makes a mistake, my teacher shows us how to learn from it. (r)	-	When someone makes a mistake, my teacher shows us how to learn from it. (r)
	When we make mistakes, our teacher gives us time to try again. (r)	"My teacher" instead of "our teacher" for consistency.	When we make mistakes, my teacher gives us time to try again. (r)
	My teacher discusses our mistakes with the whole class, so that everybody can	Delete as that increases reliability and the scale is quite long already.	-

	learn from them. (r)		
	My teacher looks annoyed or disappointed when we make mistakes.	-	My teacher looks annoyed or disappointed when we make mistakes.
	My teacher tells us off when we make a mistake.	-	My teacher tells us off when we make a mistake.
	When we make a mistake, my teacher tells us the right answer straight away.	Delete as neither classified as adaptive nor maladaptive (Tulis, 2013).	-
	When we make mistakes, my teacher stops others from making fun of it. (r)	Singular not plural.	When we make a mistake, my teacher stops others from making fun of it. (r)
	When we make a mistake in class, my teacher picks another pupil to correct it.	Delete as that increases reliability and the scale is quite long already.	-
	When we make mistakes, my teacher ignores them and quickly moves on.	Singular not plural.	When we make a mistake, my teacher ignores it and quickly moves on.

Note. R = reversed. The new teacher fixed-ability practices questionnaire items were modelled after the new student items.

Appendix D

Information and Informed Consent Forms for Parents/Guardians, Teachers, and Students

D.1 Information and opt-out form for parents/guardians

"Teachers' and their pupils' views about ability at maths and maths learning: Relationships with teaching practices"

Dear Parent/Guardian,

Your child's school has agreed to participate in a research study that involves year 7 pupils. As part of my doctoral degree at the University of Cambridge, I will be examining the relationships between the views of teachers and those of their pupils regarding ability at maths and maths learning. I will be investigating how pupils' views may be affected through different teaching practices. I am writing to tell you more about what this study will involve so you can decide if you are happy for your child to take part. If, after reading this information, you are not happy for your child to participate, please complete the opt out form. If you are happy for your child to take part, I will also ask them for their assent.

Participation is voluntary

Participation in this study is entirely voluntary. You may also withdraw your child from the study at any time without giving reasons, which will involve no penalty or loss, now or in the future.

What does participation involve?

If you are happy for your child to participate, and if your child decides to take part, (s)he will be asked to complete three questionnaires over the school year, which would take place during school hours. This should take about 20 minutes per questionnaire. Responses to most questions are given by choosing one out of five answer options. In addition, I may observe and video-record two maths lessons to learn more about the teaching practices used in the actual classroom. The teacher rather than the children will be the focus of the observations. Nevertheless, your child may be recorded as well. In addition, maths grades of your child will be obtained.

All data will be confidential and anonymised

All data will be confidential, which means that the responses of all children will be anonymised and that no individual results will be made available. The video tapes will only be available to members of the research team, which includes me, my supervisor Dr McLellan, as well as a third researcher who will help with coding. Parts of the video may be presented to academic or educational audiences but, importantly, the faces of all children will be pixelated so that no individual child can be recognised. The data collected from this study may be used for educational and publication purposes.

There are no possible disadvantages

There are no anticipated risks or discomforts for participating in this study. I hope that your child will find thinking about the questions interesting but there is no direct material benefit or financial remuneration for participating. A report about the study will be provided to the school, which may give important insights into the relationship between different teaching practices and pupils' views about ability at maths. The results may inform educators as well as policy makers about teaching methods that may foster beneficial beliefs about ability at maths in pupils.

You can contact me with any questions

If you have any questions regarding this study, please do not hesitate to contact me by email on XYZ or by phone on XYZ. You may also contact your child's Head of Maths in case that you have any concerns.

Sincerely,
Pia Kreijkes
Doctoral Candidate University of Cambridge

Parent/Guardian Opt-out form

"Teachers' and their pupils' views about ability at maths and maths learning: Relationships with teaching practices"

If you wish to opt your child out of the above described study, please fill in the slip below and ask your child to hand it to his/her maths teacher. If you are happy for your child to participate if (s)he wants so, you do not have to return this slip.

I have read the information about the study and talked about this with my child.

I am not willing for my child to take part in the study.

Please tick the box. ☐

Name of child: ....
Class: ....
Signature of parent/guardian: ....
Date: ....

D.2 Information for parents/guardians for the mixed-methods sample

"Teachers' and their pupils' views about ability at maths and maths learning:
Relationships with teaching practices"

Dear Parent/Guardian,

As you may already know, your child's school is participating in a research study with their year 7 pupils. This study is part of my doctoral degree at the University of Cambridge, which examines the relationships between teachers' and their pupils' views about ability at maths and maths learning.

Your child may have already taken part in the first part of the study, which involved questionnaires. Your child's teacher has kindly agreed to also take part in classroom observations, which involves video-recording of maths lessons. Your child may also be asked to take part in group interviews. If, after reading this information, you are not happy for your child to be recorded or interviewed, please complete the opt out form. Otherwise you do not have to return the form and your child can decide him- or herself.

Participation is voluntary

Participation in this study is entirely voluntary. You may also withdraw your child from the study at any time without giving reasons, which will involve no penalty or loss, now or in the future.

What does participation in the observation and interviews involve?

I will observe and video-record a few lessons of your child's maths class to learn more about the teaching practices that are used. The observations will be part of the usual classroom routine, which means that the lessons will not be interrupted. The teacher rather than the children will be the focus of the observations. Nevertheless, your child may be recorded as well. After observations, your child may be asked to take part in a group interview with other children from the class, which is voluntary. Interviews will explore what pupils think about different teaching practices. These interviews will be audio-recorded.

All data will be confidential and anonymised

All data will be confidential, which means that the recordings will only be available to members of the research team. Parts of the video may be presented to academic or educational audiences but, importantly, the faces of all children will be pixelated so that no individual child can be recognised. No child will be mentioned by name and only group results will be reported. The results from this study may be used for educational and publication purposes.

There are no anticipated risks

There are no anticipated risks or discomforts and there is no direct material benefit or financial remuneration for participating. A report about the study will be provided to the school and the results may inform educators and policy makers about the relationships between different teaching practices and pupils' views about ability at maths.

You can contact me with any questions

If you have any questions regarding this study, please do not hesitate to contact me by email (XYZ) or telephone (XYZ). You may also contact your child's teacher in case you have any concerns.

Sincerely,
Pia Kreijkes
Doctoral Candidate University of Cambridge

Parent/Guardian Opt-Out Form

"Teachers' and their pupils' views about ability at maths and maths learning:
Relationships with teaching practices"

If you wish to opt your child out of the above described vide-recording and/or interviews, please fill in this slip and ask your child to hand it to his/her maths teacher. If you are happy for your child to participate if he or she wants to, you do not have to return this slip.

Please tick the appropriate box/es.

☐ I am **not** willing for my child to be video-recorded.

☐ I am **not** willing for my child to take part in an interview.

Name of child: ....

Maths class: ....

Signature of parent/guardian: ....

Date: ....

D.3 Information and written informed assent form for students

Dear year 7 pupil,

My name is Pia and I'm a doctoral student at the University of Cambridge. I would like to ask you to take part in a research study. I want to find out what teachers and their pupils think about ability at maths and learning maths. I'm also interested in the different teaching methods that maths teachers use during their lessons.

It's completely up to you if you want to take part in this research. Your parent allows you to take part, but you do not have to. Even if you say 'yes' now, you can change your mind later.

What would you be asked to do?

I would ask you to fill in three questionnaires during school time (one at the beginning, one in the middle, and one at the end of the school year). These are no tests, so there are no right or wrong answers. To answer a question, you simply need to choose one out of different options. This should take about 20 minutes for each questionnaire. If you do not want to answer some of the questions, you don't have to.

I may also come to your classroom to observe what your teacher is doing during the lessons. These lessons will be video- and audio-taped. You don't have to worry about that, because I will focus on your teacher and not you. If you are interested, you may be asked to be part of a little discussion group with other children in your class. We would talk about what you think about the things your teacher is doing during maths classes. This discussion would be audio-taped. But you can decide later on if you want to do this.

Is everybody going to know about this?

We will not tell other people that you are taking part in this study. We also won't share information about you and your individual answers to anyone who doesn't work in the research team. If you are being video-recorded, I will make sure that your face cannot be recognised, so that no one will know that you are in the video. After the research is over, your school will be told what we found in the study and the results will also be shared with other people who do research or with those working in education. But no one will know about your personal answers.

Who can you talk to or ask questions to?

You can ask me questions now or later. If you want to talk to someone else that you know, for example your teacher or friends, that's okay too.

With many thanks,

Pia Kreijkes

Written Informed Assent Form for Pupils

Title of project: Teachers' and their pupils' views about ability at maths and maths learning: Relationships with teaching practices

Researcher: Pia Kreijkes
Supervisor: Dr Ros McLellan

- I understand what it means to take part in the study that the researcher and the information sheet explained to me.
- I was allowed to ask questions. I also know that I can ask questions later if I want.
- I know that my parent/guardian allows me to take part in the study but that it is my own decision if I want to or not. I can decide to stop with the study whenever I want.
- I know that I do not have to answer any questions that I do not want to answer.
- I understand that the results of the study will be shared with my school and that they may be shared with the public but that my name will never be mentioned.
- No one will know what answers I gave specifically.

I agree to take part in this study.

Print name: ___________________________

Signature: __________________________

Date: __________________________

D.4 Information and written informed consent form for teachers

“Teachers’ and their pupils’ beliefs about ability at maths and maths learning: Relationships with teaching practices”

This letter is to inform you about a study that your school has agreed to participate in, and to invite you to take part as well. As a doctoral candidate at the University of Cambridge, I am studying the relationships between teachers’ and their pupils’ beliefs about ability at maths and maths learning. This research investigates how different teaching practices relate to these beliefs as well as how pupils’ beliefs relate to their maths attainment.

What does participation involve?

If you are happy to be involved, you and your year 7 maths pupils [given that they agree to do so] will be asked to complete three questionnaires over the course of the current school year (at the beginning, in the middle, and towards the end). Pupils will do so during maths lessons, which should take no longer than 20 minutes, while you can do so online, which should take no more than 15 minutes per questionnaire.

Shortly after the first questionnaire, you may be asked to participate in an additional study part, for which participation is entirely voluntary and independent of your decision to be involved with the questionnaires. I am planning to observe a couple of maths lessons with participating classes to learn more about the things that I will be asking in the questionnaires based on the actual classroom. This would take place during a two-week period towards the middle of the school year. These lessons would be video-taped. The observed lessons should be part of the normal classroom agenda and will thus not require any additional effort on your part. After I’ve visited your lessons, I would like to talk to you to explore in more detail what your views are on different teaching practices in maths. I would also conduct group interviews with a few selected pupils of your class. These interviews will be audio-taped and should take about 30 to 45 minutes.

Participation is voluntary

Participation in this study is entirely voluntary. You may also withdraw from the study at any time without giving reasons, which will involve no penalty or loss, now or in the future.

All data will be confidential and anonymised

Any information that is obtained in connection with this study that can identify you will remain confidential. All data collected will be number-coded and any names will be removed from the data once all relevant information has been collected, so that complete anonymisation is ensured. The confidentiality of the data will be maintained throughout the time it is required to be stored after which time it will be destroyed. The data collected from this study will be used for educational and publication purposes. Group rather than individual results will be presented whenever possible and pseudonyms will be used throughout any reports.

There are no possible disadvantages

There are no anticipated risks or discomforts for participating in this study and ethical approval has been granted by the University of Cambridge.

As a participant, you will not receive a direct material benefit or financial remuneration.

Nevertheless, your school will receive a report of the findings specific to your own school. This will provide you with interesting insights into the beliefs that teachers and pupils in your school hold regarding ability at maths and maths learning. The report would also inform you about the teaching practices that are commonly employed during maths lessons at your school and provide you with a discussion of these.

You can contact me with any questions

If you have any questions regarding participating in this study, please do not hesitate to contact me by email on XYZ or by phone on XYZ.

This study can only be successful with the help of schools, teachers and pupils. I thus hope that you decide to participate.

Sincerely,

Pia Kreijkes
Doctoral Candidate
University of Cambridge

Informed Consent Form for Teachers

Title of project: Teachers' and their pupils' beliefs about ability at maths and maths learning: Relationships with teaching practices

Researcher: Pia Kreijkes
Supervisor: Dr Ros McLellan

I have understood the details of the research as explained to me by the information letter and the researcher. I have had an opportunity to ask questions and have them answered. I understand that my participation is entirely voluntary and that the data collected during the study will not be identifiable. Moreover, I am aware that I have the right to withdraw from the study at any time without providing any reasons for doing so. I understand that the data I provide may be used for analysis and publication after being completely anonymised so that I cannot be identified. Finally, I know that I can contact Pia Kreijkes (XYZ, XYZ) at the Faculty of Education if I have any questions.

I hereby confirm that I agree to take part in this study.

Sign name: ___________________________

Print name ____________________________

Date: ___________________________

Email address (used to send you the link to the online questionnaires):

____

Appendix E
Example Notes of Classroom Observation

T = Teacher
S = Student

Table E.1

Example Notes of Classroom Observation

Time	Notes
	T has to pick the class up after lunch break. Thus classroom is empty but the task is already on the smart board. T already distributed a marked test from a previous lesson. Two Ss are already coming to the room and check their results. They also walk around to see what other Ss got. (public evaluation?)
1.25	Class enters, Ss get handshake. They go to their places and start working on the task on the board. Topic is: Recognise fractional fact families. T tells Ss to put the parent letter on the table if they will not take part in the study. Says that they do not have to say anything. No S does it so I turn on the camera. T talks to S. S appears to say that he does not want to be filmed so T asks him to swap with a S on the back of the room. I turned cam off for this. S says something about her glue stick and T says she should borrow it. S says something. T: Oh it's on the floor. S in front of her picks it up and hands it over.
1.30	Alarm rings. T: In 5 seconds, stick in your test….A bit longer because some are going very slow. T starts to discuss work on the board. Finding fact families. T asks Ss to give her the facts for different equations on the board. T praises S with: Excellent. (product-praise) Next equation… T: What is 6*5? S: 30. T: How else can we express this relationship? S: 30:6=5. T gives praise. T: What if I wanted to use a fraction there? S: 6/30. T: What's the simplest form. S: 1/5. . T: Can you give me another multiplication from the same fact family. S says something. T: Fantastic. One more fact family. I give you a hint because I want you to be confident. (risk-taking?) T writes a fact family on board as an example. T: What's my missing fact then? S: 30:1/6. (mistake)

	T explains what she did in the example. S makes mistake. T: No, it will be a division. (reaction to mistake) S: 5:1/6=30. T: Who did divisions with fractions in primary school? A few Ss put their hands up. T: Good, you should have been bolder. T explains how to express this equation (dividing by a fraction). T:…just skipped four lessons because I got overexcited. (high expectations?) T shows a new bar model on board. T: Nothing in your hands. Sit straight. T: What is ½*40 equivalent to? S: 40:2. T: Fantastic. T: And what is 1/3*30 equivalent to. S? S: 30:3. T: Excellent. S says something about this. T: Exactly….
1.41	T puts new task up. T: This is complex so you need to listen carefully. T explains the task. T: Two minutes to talk with person next to you. I will call you randomly, so make sure you have an answer. (partner work and presumably most participate) Ss start discussing. T: No fraction needed guys. No fractions needed. T talks to S.
1.45	Alarm rings. (All very structured - Set times for everything) T: 5…4…3…2…1. Tell me what goes in the pink boxes. S? S gives correct answer. T repeats it and says: Good. What about the orange. S gives correct answer. T: Fantastic. T repeats it. (Seems to be mostly praise for correct answers) . S says something for one of the tasks about division but there is already a multiplication sign in the task. T: Thanks for pointing that out, X. It's very helpful. T explains it further. T puts up new slide showing representations of the same number. (Multidimension tasks) T: What's the same and what's different between these representations. 1 Minute with person next to you. Ss talk about it. T: What's the same? S: It's half of 8. T explains it. T: How do I know its ½ of 8? S explains it. T: Good, excellent. What's different? S: They present them in different ways…. T: Excellent…

Appendix F
Instructions for Questionnaire Administration

Please explain the following:

- The questions are about what students think about learning maths and what is important for them in their maths lessons.
- The questions might remind them of the first questionnaire from the beginning of the year, but they should just think about what their opinions are now.
- It works exactly like the last questionnaire, so they simply have to tick boxes.
    - It is not possible to tick more than one box for a question and they have to **decide for one answer** (i.e., no middle options or two options allowed).
    - They can skip questions if they really do not want to answer it but I hope they will answer all.
- It is very important that they give their **honest** opinion.
- **Their individual answers will not be shared with anyone, including teachers.**
    - Please make it clear that you will not look at their answers (and, obviously, please don't look at their answers).
- They should **provide their full name** on the first page but their name will not be shared with others - this is only to match up their answers from the other two questionnaires.
- They should also **provide your name (their current main teacher)** on the first page.
- If they struggle to understand a question, they are allowed to ask for clarifications.
- Some of the questions are very similar but they should not bother about that and rather take one question after the other.

Appendix G

Results for Exploratory Structural Equation Modelling

Table G.1

Results of the First-Order Confirmatory ESEM: Factor Loadings and Standardised Factor Correlations

Factor	Item ID	1	2	3	4	5
1. Unsophisticated task differentiation	1. TD1	**.90**				
	2. TD2	**.63**				
	3. TD3	**.57**				
	4. TD4	**.73**				
2. Public evaluation	20. EP2		**.50**			
	22. EP4		**.54**			
	24. ES1		**.81**			
	25. ES2		**.71**			
	27. ES4		**.65**			
3. Promotion of self-regulation	34. ASR2			**.38**		
	35. ASR3			**.38**		
	36. ASR4			**.61**		
	37. ASR5			**.99**		
4. Differential value	38. DV1				**.63**	
	39. DV2				**.49**	
	40. DV3				**.80**	
	41. DV4				**.82**	
	42. DV5				**.75**	
5. Promotion of risk-taking	19. EP1					**.44**
	44. RH2					**.64**
	47. RM1					**.80**
	48. RM2					**.58**
	52. RM6					**.58**
Standardised factor correlations						
1. Unsophisticated task differentiation		-				
2. Public evaluation		**.31**	-			
3. Promotion of self-regulation		.07	.06	-		
4. Differential value		**.40**	**.45**	**-.33**	-	
5. Promotion of risk-taking		-.10	-.13	**.56**	**-.61**	-

Note. N = 420. Loadings < .32 are suppressed. χ^2 (148) = 177.16, p = .05, CFI = 0.993, TLI = 0.988, RMSEA = 0.022 [<0.001; 0.033], SRMR = 0.021. Factors 3 and 5 reflect adaptive practices. All primary item loadings (in bold) were statistically significant at $p < .01$.

Table G.2

Results of the Hierarchical Confirmatory ESEM: Factor Loadings and Standardised Higher-Order Uniform Factor Loadings

Factor	Item ID	1	2	3	4	5
1. Unsophisticated task differentiation	1. TD1	**.90**		.44		
	2. TD2	**.66**			**.40**	
	3. TD3	**.56**		.33		
	4. TD4	**.73**			.42	
2. Public evaluation	20. EP2		**.55**			
	22. EP4		**.56**			
	24. ES1		**.81**			
	25. ES2		**.71**			
	27. ES4		**.63**			
3. Promotion of self-regulation	34. ASR2			**.43**		**.34**
	35. ASR3			**.40**		**.44**
	36. ASR4			**.61**		**.52**
	37. ASR5			**.99**		**.40**
4. Differential value	38. DV1				**.64**	
	39. DV2				**.47**	
	40. DV3				**.80**	
	41. DV4				**.82**	
	42. DV5				**.75**	
5. Promotion of risk-taking	19. EP1					**.51**
	44. RH2					**.64**
	47. RM1					**.80**
	48. RM2					**.61**
	52. RM6					**.50**
Standardised higher-order uniform factor loadings						
Higher-order factor		.03	.14	-.01	3.55	-.18

Note. N = 420. Loadings < .32 are suppressed. χ^2 (153) = 181.14, p = .06, CFI = 0.993, TLI = 0.988, RMSEA = 0.021 [<0.001; 0.032], SRMR = 0.022. Factors 3 and 5 reflect positive practices. Loadings in bold were statistically significant at $p < .01$.

Table G.3

Results of the Bifactor ESEM: Factor Loadings

Factor	Item ID	1	2	3	4	5	G
1. Unsophisticated task differentiation	1. TD1	**.82**					**.21**
	2. TD2	**.56**					**.49**
	3. TD3	**.57**					.17
	4. TD4	**.69**					**.27**
2. Public evaluation	20. EP2		**.48**				.15
	22. EP4		**.53**				-.01
	24. ES1		**.71**				**.46**
	25. ES2		**.62**				**.31**
	27. ES4		**.60**				**.40**
3. Promotion of self-regulation	34. ASR2			**.35**			**-.40**
	35. ASR3			**.35**			**-.45**
	36. ASR4			**.56**			**-.33**
	37. ASR5			**.83**			**-.31**
4. Differential value	38. DV1				.14		**.76**
	39. DV2				**.36**		**.51**
	40. DV3				.33		**.75**
	41. DV4				.24		**.89**
	42. DV5				.07		**.91**
5. Promotion of risk-taking	19. EP1					**.29**	**-.53**
	44. RH2					**.44**	**-.59**
	47. RM1					**.60**	**-.53**
	48. RM2					**.31**	**-.58**
	52. RM6					**.26**	**-.62**

Note. N = 420. Cross-loadings < .32 are suppressed. χ^2 (130) = 151.59, p = .09, CFI = 0.995, TLI = 0.990, RMSEA = 0.020 [<0.001; 0.032], SRMR = 0.019. Factors 3 and 5 reflect positive practices. Loadings in bold are statistically significant at *p* < .01.

Table G.4

Results of the First-Order Confirmatory ESEM with the Entire Sample: Factor Loadings and Standardised Factor Correlations

Factor	Item ID	1	2	3	4	5
1. Unsophisticated task differentiation	1. TD1	**.87**				
	2. TD2	**.68**				
	3. TD3	**.56**				
	4. TD4	**.75**				
2. Public evaluation	20. EP2		**.50**			
	22. EP4		**.58**			
	24. ES1		**.73**			
	25. ES2		**.62**			
	27. ES4		**.65**			
3. Promotion of self-regulation	34. ASR2			**.38**		
	35. ASR3			**.47**		
	36. ASR4			**.59**		
	37. ASR5			**.91**		
4. Differential value	38. DV1				**.65**	
	39. DV2				**.48**	
	40. DV3				**.79**	
	41. DV4				**.87**	
	42. DV5				**.73**	
5. Promotion of risk-taking	19. EP1					**.42**
	44. RH2					**.73**
	47. RM1					**.87**
	48. RM2					**.57**
	52. RM6					**.57**
Standardised factor correlations						
1. Unsophisticated task differentiation		-				
2. Public evaluation		**.32**	-			
3. Promotion of self-regulation		.08	.11	-		
4. Differential value		**.40**	**.40**	**-.33**	-	
5. Promotion of risk-taking		**-.11**	-.03	**.57**	**-.63**	-

Note. N = 831. Loadings < .32 are suppressed. χ^2 (148) = 195.99, p = .01, CFI = 0.993, TLI = 0.987, RMSEA = 0.02 [0.011; 0.027], SRMR = 0.018. Factors 3 and 5 reflect adaptive practices. All primary item loadings (in bold) were statistically significant at p < .01.

Appendix H
Technical Details for Regression Models

H.1 Irregularities in assigning weights for group membership

Some irregularities regarding the weighting scheme must be noted. In the first school, one teacher who taught the lowest ability set was replaced after the first term and no data about her teaching practices were obtained. Accordingly, her teaching practices were not taken into account. This affected nine students who have been taught by her for 18 weeks and nine students who have been taught by her for 9 weeks. In School 2, two classes were completely excluded from the analyses because the teachers changed multiple times and teaching practice data were only obtained for a teacher that taught the class for only six weeks. Eight students who belonged to these classes during the first term and then changed classes were included in analyses, taking only the teaching practices of the second teacher into account. For three other classes, the second motivational beliefs questionnaire was administered a couple of months earlier than in the other classes due to a teacher change. Hence, potential influences of the subsequent teacher on achievement were not taken into account.

In School 3, one of the two participating classes was taught by a teacher who did not participate, and for whom no teaching data was available, for three out of seven classes. Hence, only the other teacher was considered. In School 4, one class was excluded because the teacher changed after the first term and no teaching data was obtained for the following three teachers. One class was taught by two teachers throughout the year but only the main teacher (four out of five classes) was taken into account because no more data could be obtained. In School 5, one teacher was replaced by another after 18 weeks of teaching, and only teaching practices for the subsequent teacher were taken into account. In School 6, one class was taught by two teachers simultaneously during the first 25 weeks of school and one of these was then replaced by another teacher for the last 18 weeks. Students only reported practices for the teacher that left and the one that replaced the teacher but not for the teacher who stayed. However, another class did report practices for that teacher, their main teacher, so that their classroom aggregated FAPs score was also used for the other class. These scores were also used for a third class that was taught by this teacher, in which students only reported practices for their other two teachers. While this is not ideal because practices of the same teacher might differ between classes, it was the most feasible option given that it would have been too large of a burden to ask students to report on multiple teachers.

H.2 Formulae for regression models

This Appendix presents all regression formulae described in Section 6.3. The following models were estimated for every outcome variable, namely, fixed mindset, mastery goal orientation, performance-approach goal orientation, performance-avoidance goal orientation, failure-to-ability attribution, self-concept, and achievement:

The two-level cross-classified random intercept model for the single-teacher sample:

The formula at Level 1 is

$Y_{i(t,c)} = \beta_{0(t,c)} + e_{i(t,c)}$, (1)

and, at Level 2

$\beta_{0(t,c)} = \gamma_{000} + u^1_{0t0} + u^2_{00c}$, (2)

where,

- $Y_{i(t,c)}$ is the outcome for student i who belongs to teacher t and classroom c
- $\beta_{0(t,c)}$ is the average outcome for teacher t and classroom c
- $e_{i(t,c)}$ is the student-level residual associated with student i who belongs to teacher t and classroom c, with $e_{i\{t\}\{c\}} \sim \mathrm{SkewNormal}(0, \omega 2, \alpha)$
- γ_{000} is the outcome score averaged across all teachers and classrooms
- u^1_{0t0} and u^2_{00c} are the second-level residuals or the variability around the intercept, γ_{000} associated with teacher t and classroom c, respectively.

The two-level cross-classified multiple membership random intercept model for the entire sample:

The formula at Level 1 is

$Y_{i(\{h\},\{g\})} = \beta_{0(\{h\},\{g\})} + e_{i(\{h\},\{g\})}$, (3)

and, at Level 2

$\beta_{0(\{h\},\{g\})} = \gamma_{000} + \sum_{t \in \{h\}} w_{t0} u^1_{0t0} + \sum_{c \in \{g\}} w_{i0c} u^2_{00c}$, (4)

where,

- $Y_{i(\{h\},\{g\})}$ is the outcome for student i who is a member of a set of teachers {h} and classrooms {g}
- $\beta_{0(\{h\},\{g\})}$ is the average outcome for the set of teachers {h} and classrooms {g}
- $e_{i(\{h\},\{g\})}$ is the student-level residual associated with student i, with $e_{i(\{h\},\{g\})} \sim \mathrm{SkewNormal}(0, \omega 2, \alpha)$
- γ_{000} is the outcome averaged across all sets of teachers and classrooms
- w_{it0} is the predetermined weight assigned to student i's association with teacher t in set h, where $\sum_{t \in \{h\}} w_{it0}$ adds to 1
- w_{i0c} is the predetermined weight assigned to student i's association with classroom c in set g, where $\sum_{c \in \{g\}} w_{i0c}$ adds to 1
- u^1_{0t0} is the residual associated with teacher t or the variance of the mean for teacher t around the overall mean, with $u^1_{0t0} \sim \mathrm{SkewNormal}(0, \omega 2, \alpha)$
- u^2_{00c} is the residual associated with classroom c or the variance of the mean for classroom c around the overall mean, with $u^2_{00c} \sim \mathrm{SkewNormal}(0, \omega 2, \alpha)$

The two-level cross-classified random effects model for the single-teacher sample: This model includes sex (girl), the lagged outcome, and the five FAPs dimensions at the student level and the uniform FAPs dimension at the teacher/classroom level. All student-level effects were modelled as random. The control variables and the aggregated FAPs score were grand-mean centred (CGM) and the five student-level FAPs dimensions were centred within the cluster (CWC).

The formula at Level 1 is

$$Y_{i(t,c)} = \beta_{0(t,c)} + \beta_{1(t,c)}\boldsymbol{Sex\_CGM}_{i(t,c)} + \beta_{2(t,c)}\boldsymbol{Ylagged\_CGM}_{i(t,c)} + \beta_{3(t,c)}\boldsymbol{TaskDif\_CWC}_{i(t,c)} + \beta_{4(t,c)}\boldsymbol{PubEva\_CWC}_{i(t,c)} + \beta_{5(t,c)}\boldsymbol{SelfReg\_CWC}_{i(t,c)} + \beta_{6(t,c)}\boldsymbol{DifValue\_CWC}_{i(t,c)} + \beta_{7(t,c)}\boldsymbol{Risk\_CWC}_{i(t,c)} + e_{i(t,c)}, \quad (5)$$

and, at Level 2

$$\beta_{0(t,c)} = \gamma_{000} + u^1_{0t0} + u^2_{00c} + \gamma_{011}\boldsymbol{FAPs\_CGM}_{tc} + u^1_{011t0} + u^2_{0110c},$$

$$\beta_{1(t,c)} = \gamma_{100} + u^1_{1t0} + u^2_{10c},$$

$$\beta_{2(t,c)} = \gamma_{200} + u^1_{2t0} + u^2_{20c},$$

$$\beta_{3(t,c)} = \gamma_{300} + u^1_{3t0} + u^2_{30c},$$

$\beta_{4(t,c)} = \gamma_{400} + u^1_{4t0} + u^2_{40c}$,

$\beta_{5(t,c)} = \gamma_{500} + u^1_{5t0} + u^2_{50c}$,

$\beta_{6(t,c)} = \gamma_{600} + u^1_{6t0} + u^2_{60c}$,

$\beta_{7(t,c)} = \gamma_{700} + u^1_{7t0} + u^2_{70c}$, (6)

Where, in addition to or other than formulas 1 and 2,

- $\beta_{0(t,c)}$ represents the average outcome for teacher t and classroom c when the control variables and the aggregated FAPs score are at the grand-mean and the student-level predictors are at the cluster mean
- γ_{000} represents the average outcome when the control variables and the aggregated FAPs score are at the grand-mean and the student-level predictors are at the cluster mean
- $\beta_{1(t,c)}$ to $\beta_{7(t,c)}$ represent the change in the outcome for teacher t and classroom c per unit change in the respective predictor while all other predictors are held constant
- γ_{100} to γ_{700}represents the average change in the outcome per unit change in the respective predictor while all other predictors are held constant
- γ_{011} represents the average change in the intercept per unit change in $\boldsymbol{FAPs\_CGM}_{tc}$
- u^1_{0t0} and u^2_{00c} represent the unexplained teacher and class-level variance for teacher t and classroom c, respectively, around the average outcome after controlling for all predictors in the model
- u^1_{1t0} to u^1_{011t0} and u^2_{10c} to u^2_{0110c} represent the unexplained teacher and class-level variance around the average slope after controlling for all predictors in the model
- $e_{i(\{h\},\{g\})}$ is the student-level residual associated with student i after controlling for all predictor variables, with $e_{i(\{h\},\{g\})} \sim \text{SkewNormal}(0, \omega 2, \alpha)$

The two-level cross-classified multiple membership random effects model: This model includes sex (girl) and the lagged outcome at the student level, which were modelled as random effects, and the aggregated FAPs score at the second level. All predictor variables have been grand-mean centred (CGM).

The formula at Level 1 is

$Y_{i(\{h\},\{g\})} = \beta_{0(\{h\},\{g\})} + \beta_{1(\{h\},\{g\})}\boldsymbol{Sex\_CGM}_{i(\{h\},\{g\})} + \beta_{2(\{h\},\{g\})}\boldsymbol{Ylagged\_CGM}_{i(\{h\},\{g\})} + e_{i(\{h\},\{g\})}$, (7)

and, at Level 2

$\beta_{0(\{h\},\{g\})} = \gamma_{000} + \sum_{t\in\{h\}} w_{it0}(\gamma_{011}\boldsymbol{FAPs\_CGM_{tc}} + u^1_{0t0}) + \sum_{c\in\{g\}} w_{i0c}(\gamma_{011}\boldsymbol{FAPs\_CGM_{tc}} + u^2_{00c})$,
$\beta_{1(\{h\},\{g\})} = \gamma_{100} + \sum_{t\in\{h\}} w_{it0}u^1_{1t0} + \sum_{c\in\{g\}} w_{i0c}u^2_{10c}$,
$\beta_{2(\{h\},\{g\})} = \gamma_{200} + \sum_{t\in\{h\}} w_{it0}u^1_{2t0} + \sum_{c\in\{g\}} w_{i0c}u^2_{20c}$, (8)

where, in addition to or other than formulas 3 and 4,

- $\beta_{0(\{h\},\{g\})}$ is the average outcome for the set of teachers h and classrooms g when all predictors are at the grand-mean
- $\beta_{1(\{h\},\{g\})}$and $\beta_{2(\{h\},\{g\})}$represent the average increase in the outcome for the sets of teacher h and classes g while all other variables are held constant
- $e_{i(\{h\},\{g\})}$ is the student-level residual associated with student i after controlling for all predictors, with $e_{i(\{h\},\{g\})} \sim \mathrm{SkewNormal}(0, \omega2, \alpha)$
- γ_{000} is the average outcome when all predictor variables are at the grand mean
- γ_{011}represents the change in the intercept per unit change in $\mathrm{FAPs\_CGM_{tc}}$ while all other values in the model are held constant
- γ_{100} and γ_{200} represent the average change in the outcome per unit increase in the respective predictor when all other predictors are held constant
- u^1_{0t0} is the residual associated with teacher t or the variance of the mean for teacher t around the overall mean after controlling for all predictors, with $u^1_{0t0} \sim \mathrm{SkewNormal}(0, \omega2, \alpha)$
- u^2_{00c} is the residual associated with classroom c or the variance of the mean for classroom c around the overall mean after controlling for all predictors, with $u^2_{00c} \sim \mathrm{SkewNormal}(0, \omega2, \alpha)$
- u^1_{1t0} and u^1_{2t0} as well as u^2_{10c} and u^2_{20c} represent the variance around the average slope for the set of teachers h and classrooms g, respectively, after controlling for all predictors

H.3 Priors for Bayesian regression analyses

Table H.1

Priors Used in the Bayesian Regression Models

Parameter class	Coefficient/Predictor (Time 1)	Outcome (Time 3)	Brms default prior	Custom prior
Population-level fixed effects (regression coefficients)	Intercept	Fixed mindset, mastery goal, performance-approach goal, performance-avoidance goal, self-concept	Improper flat prior over the reals	N(2,2)
	Intercept	Achievement	Improper flat prior over the reals	N(0,1)
	Threshold	Failure-to-ability attribution	Half-Student-t(3, 0, 10)	-
	Sex, lagged outcome, unsophisticated task differentiation, public evaluation, little promotion of self-regulation, differential value, little promotion of risk-taking, aggregated FAPs	Fixed mindset, mastery goal, performance-approach goal, performance-avoidance goal, self-concept	Improper flat prior over the reals	N(0,2)
	Sex, lagged outcome, unsophisticated task differentiation, public evaluation, little promotion of self-regulation, differential value, little promotion of risk-taking, aggregated FAPs	Achievement, failure-to-ability attribution	Improper flat prior over the reals	N(0,1)
Standard deviations of group-level (random) effects	-	Fixed mindset, mastery goal, performance-approach goal, performance-avoidance goal, self-concept, achievement	Half-Student-t(3, 0, 10)	-
Correlations of group-level ('random') effects	-	Fixed mindset, mastery goal, performance-approach goal, performance-avoidance goal, self-concept, achievement	LKJ Cholesky prior with a shape parameter of 1 (this prior gives equal likelihoods to all correlation matrices).	-
Distance parameters (simplex) of monotonic effects	Failure-to-ability attribution	Failure-to-ability attribution	Dirichlet prior with α = 1 (this assumes that all differences between adjacent categories are	-

			the same on average but allows considerable uncertainty)	
Alpha (Skew for the outcome variable)	-	Fixed mindset, mastery goal, performance-approach goal, performance-avoidance goal, self-concept, achievement	N(0, 4)	-
Sigma (Standard deviation for the outcome variable)	-	Fixed mindset, mastery goal, performance-approach goal, performance-avoidance goal, self-concept, achievement	Half-Student-t(3, 0, 10)	-

Appendix I

Results for the Longitudinal Measurement Invariance Testing

Table I.1

Results of the Unconstrained Configural ESEM for Motivational Beliefs: Standardised Loadings and Factor Correlations for Time 1

Factor	Item ID	1	2	3	4	5
1. Fixed mindset	17. P1F1	**.66**				
	15. P1F2	**.74**				
	26. P1F3	**.69**				
	07. P1F4	**.63**				
2. Mastery goal	22. P1M1		**.82**			
	02. P1M2		**.69**			
	14. P1M3		**.79**			
	19. P1M4		**.51**			
	09. P1M5		**.64**			
3. Performance-approach goal	01. P1Ap1			**.37**	**.45**	
	06. P1Ap2			**.48**		
	12. P1Ap3			**.71**		
	16. P1Ap4			**.88**		
	23. P1Ap5			**.82**		
4. Performance-avoidance goal	04. P1Av1				**.59**	
	10. P1Av2				**.51**	
	13. P1Av3				**.51**	
	18. P1Av4				**.50**	
5. Self-concept	28. P1SC1					**.58**
	29. P1SC2					**.67**
	30. P1SC3					**.93**
	31. P1SC4					**.92**
	32. P1SC5					**.75**
Standardised factor correlations						
1. Fixed mindset		-				
2. Mastery goal		**-.38**	-			
3. Performance-approach goal		**.31**	**.14**	-		
4. Performance-avoidance goal		**.21**	**.10**	**.65**	-	
5. Self-concept		**-.28**	**.35**	**.18**	*.06*	-

Note. N = 759. Loadings < .32 are suppressed. Model fit for the model including motivational beliefs at Time 1 and 3: χ^2 (777) = 927.68, p = <.001, CFI = 0.988, TLI = 0.984, RMSEA = 0.016 [0.012; 0.020], SRMR = 0.024. Loadings in bold are statistically significant at $p < .01$. Italicised loadings are significant at $p < .05$.

Table I.2

Results of the Unconstrained Configural ESEM for Motivational Beliefs: Standardised Loadings and Factor Correlations for Time 3

Factor	Item ID	1	2	3	4	5
1. Fixed mindset	17. P3F1	**.59**				
	15. P3F2	**.75**				
	26. P3F3	**.57**				
	07. P3F4	**.71**				
2. Mastery goal	22. P3M1		**.84**			
	02. P3M2		**.70**			
	14. P3M3		**.80**			
	19. P3M4		**.46**			
	09. P3M5		**.81**			
3. Performance-approach goal	01. P3Ap1			**.46**		
	06. P3Ap2			**.61**		
	12. P3Ap3			**.71**		
	16. P3Ap4			**.93**		
	23. P3Ap5			**.84**		
4. Performance-avoidance goal	04. P3Av1				**.82**	
	10. P3Av2				**.43**	
	13. P3Av3			**.30**	**.37**	
	18. P3Av4			**.32**	**.46**	
5. Self-concept	28. P3SC1					**.53**
	29. P3SC2					**.69**
	30. P3SC3					**.96**
	31. P3SC4					**.85**
	32. P3SC5					**.84**
Standardised factor correlations						
1. Fixed mindset		-				
2. Mastery goal		**-.49**	-			
3. Performance-approach goal		**.22**	*.09*	-		
4. Performance-avoidance goal		**.26**	*.03*	**.67**	-	
5. Self-concept		**-.39**	**.36**	**.16**	*.02*	-

Note. N = 759. Loadings < .30 are suppressed. Model fit for the model including motivational beliefs at Time 1 and 3: χ^2 (777) = 927.68, p = <.001, CFI = 0.988, TLI = 0.984, RMSEA = 0.016 [0.012; 0.020], SRMR = 0.024. Loadings in bold are statistically significant at $p < .01$. Italicised loadings are significant at $p < .05$.

Table I.3

Standardised Factor Correlations for Motivational Beliefs Across Time Points

	1. T3	2. T3	3. T3	4. T3	5. T3
1. Fixed mindset T1	**.56**	**-.34**	**.13**	**.16**	**-.22**
2. Mastery goal T1	**-.31**	**.57**	**.14**	.09	**.29**
3. Performance-approach goal T1	**.14**	>-.001	**.50**	**.37**	**.18**
4. Performance-avoidance goal T1	**.16**	-.06	**.40**	**.58**	.05
5. Self-concept T1	**-.27**	**.23**	**.12**	*.08*	**.71**

Note. N = 759. T = Time. Correlations in bold are statistically significant at $p < .01$. Italicised loadings are significant at $p < .05$.

Appendix J
Additional Results for the CC and CCMM Models

J.1 Posterior predictive checks (PPC) for the final CC regression model

This appendix presents the PPC plots for the two-level cross-classified random effects model predicting students' motivational beliefs and achievement in the single-teacher sample. Please note that the Kernel density plots show y (outcome) values below 0 and above 4 for variables that have a scale from 0 to 4. This is due to smoothing and is thus a visual effect.

Figure J.1
Kernel Density Plot for Fixed Mindset in the CC Model

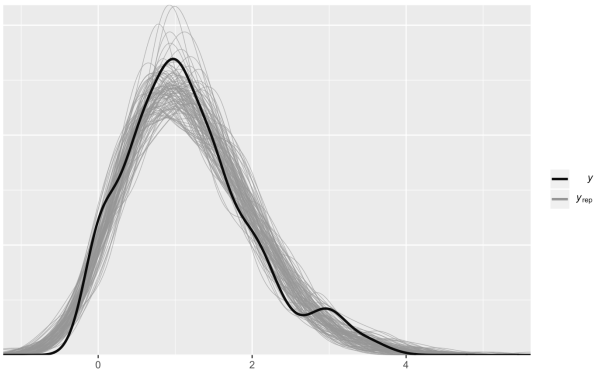


Figure J.2
Boxplot for Fixed Mindset in the CC Model

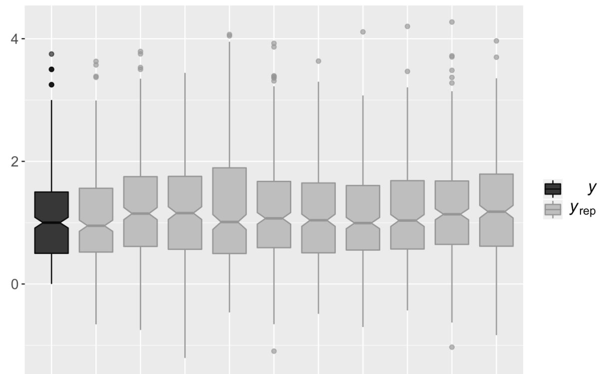


Figure J.3
Kernel Density Plot for Mastery Goal in the CC Model

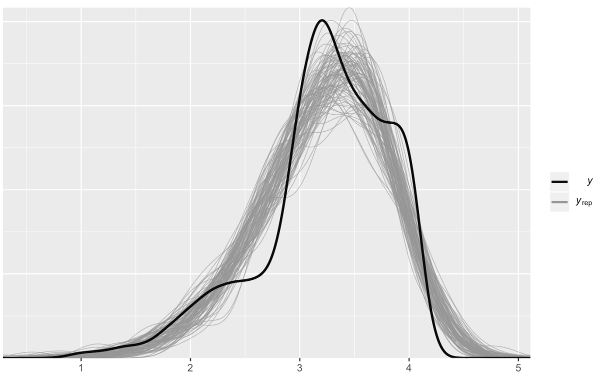


Figure J.4
Boxplot for Mastery Goal in the CC Model

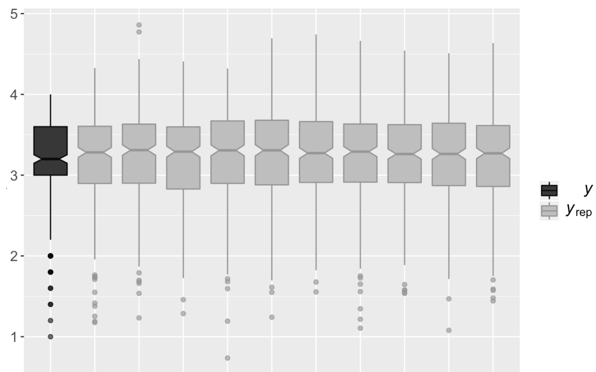


Figure J.5
Kernel Density Plot for Performance-Approach Goal in the CC Model

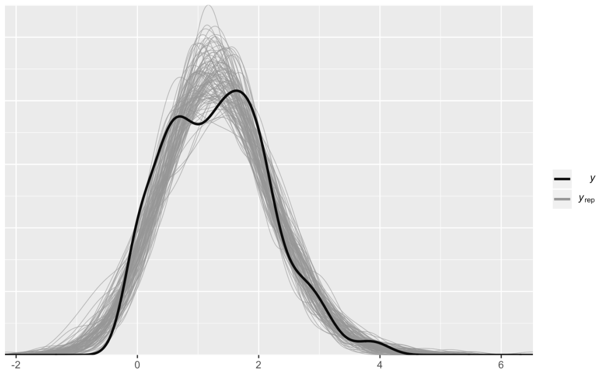


Figure J.6
Boxplot for Performance-Approach Goal in the CC Model

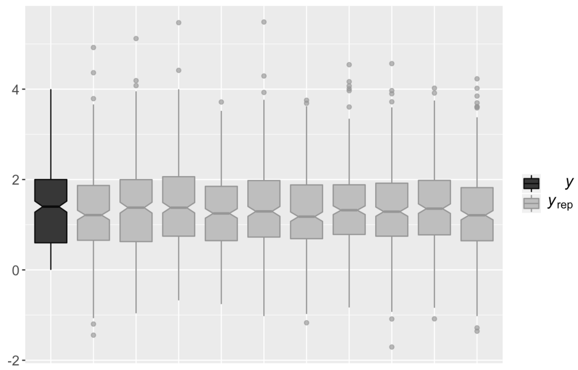


Figure J.7
Kernel Density Plot for Performance-Avoidance Goal in the CC Model

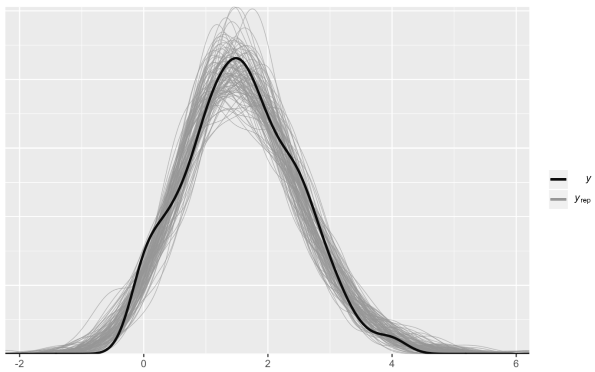


Figure J.8
Boxplot for Performance-Avoidance Goal in the CC Model

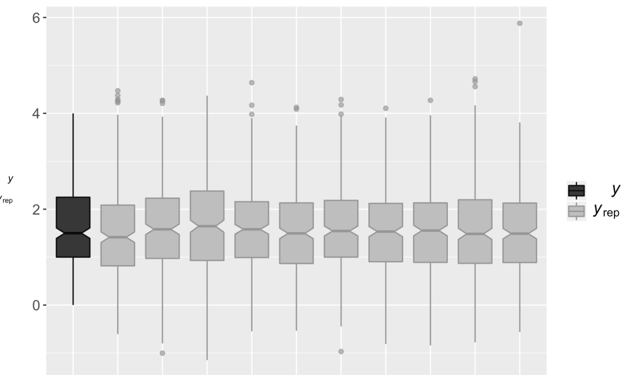


Figure J.9
Kernel Density Plot for Self-Concept in the CC Model

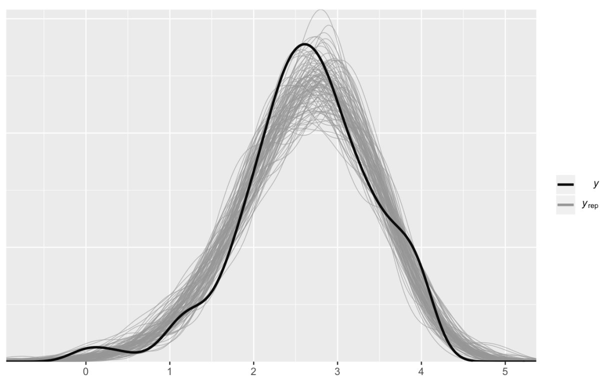


Figure J.10
Boxplot for Self-Concept in the CC Model

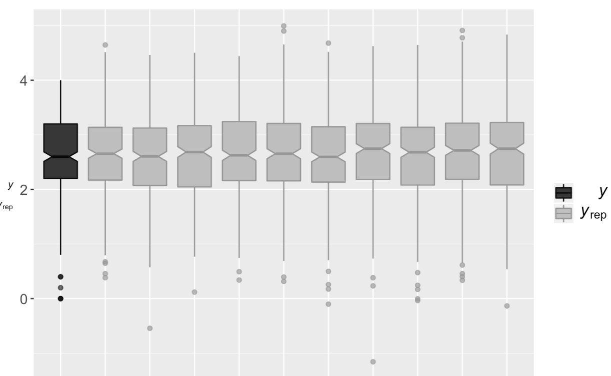


Figure J.11
Kernel Density Plot for Achievement in the CC Model

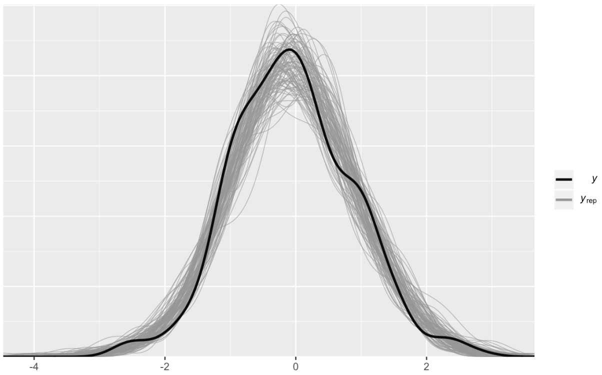


Figure J.12
Boxplot for Achievement in the CC Model

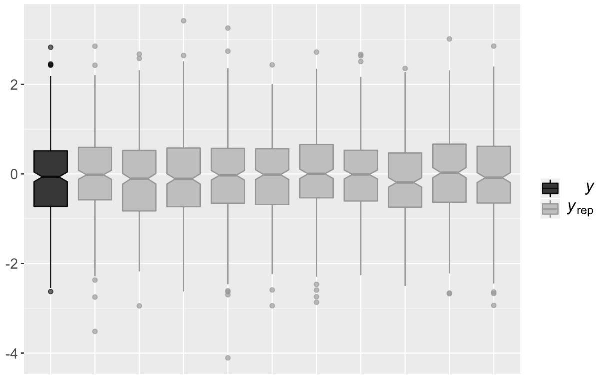
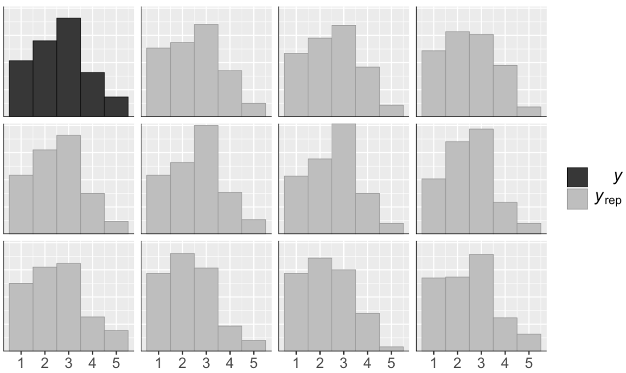


Figure J.13
Histogram for Failure-to-Ability Attribution in the CC Model



J.2 Posterior predictive checks (PPC) for the final CCMM regression model

This appendix presents the PPC plots for the two-level cross-classified multiple membership random effects model predicting students' motivational beliefs and achievement using the entire sample. Again, please note that the Kernel density plots show y values below 0 and above 4 for variables that have a scale from 0 to 4, which is a visual effect.

Figure J.14
Kernel Density Plot for Fixed Mindset in the CCMM Model

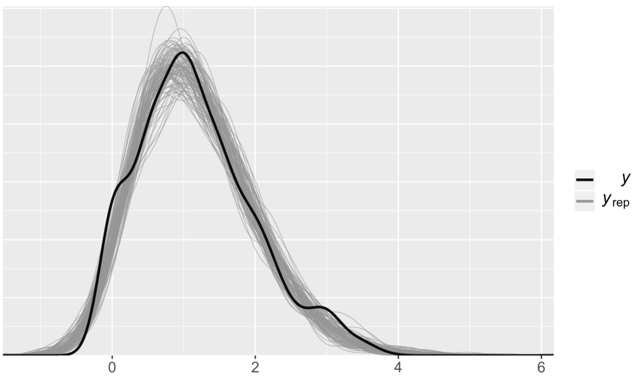


Figure J.15
Boxplot for Fixed Mindset in the CCMM Model

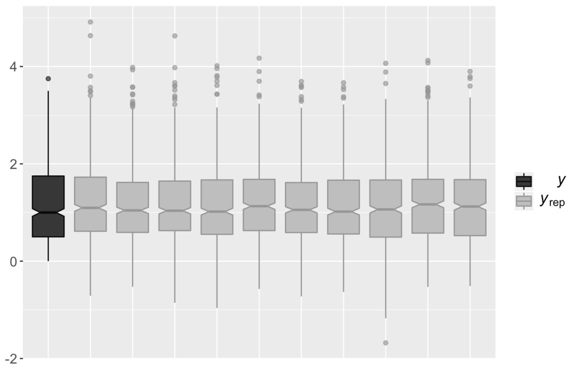


Figure J.16
Kernel Density Plot for Mastery Goal in the CCMM Model

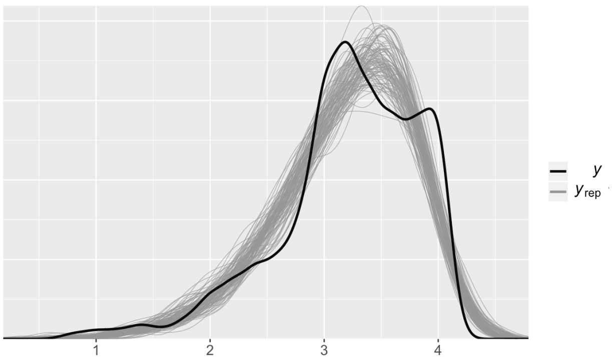


Figure J.17
Boxplot for Mastery Goal in the CCMM Model

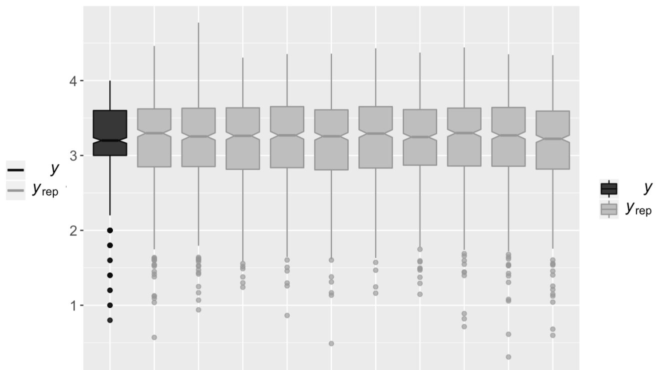


Figure J.18
Kernel Density Plot for Performance-Approach Goal in the CCMM Model

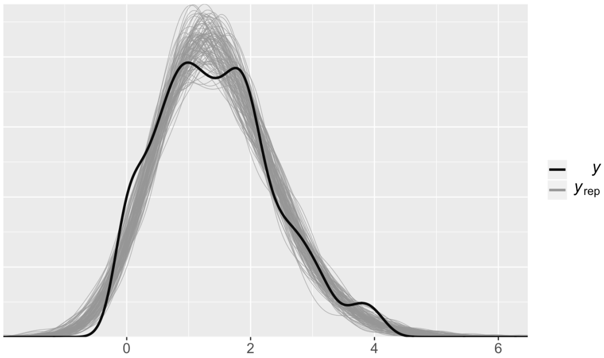


Figure J.19
Boxplot for Performance-Approach Goal in the CCMM Model

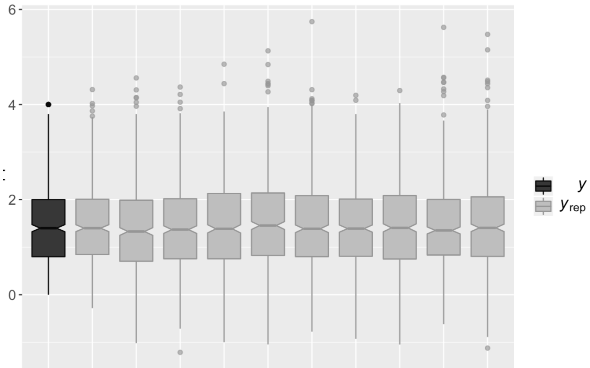


Figure J.20
Kernel Density Plot for Performance-Avoidance Goal in the CCMM Model

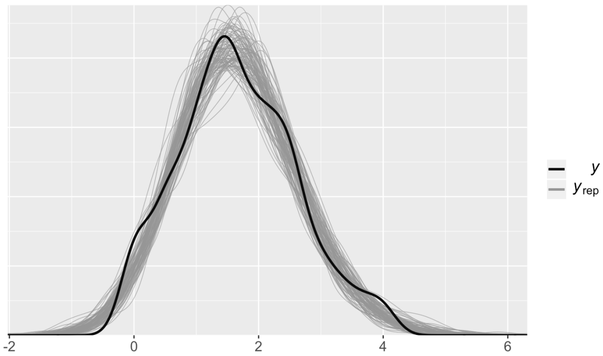


Figure J.21
Boxplot for Performance-Avoidance Goal in the CCMM Model

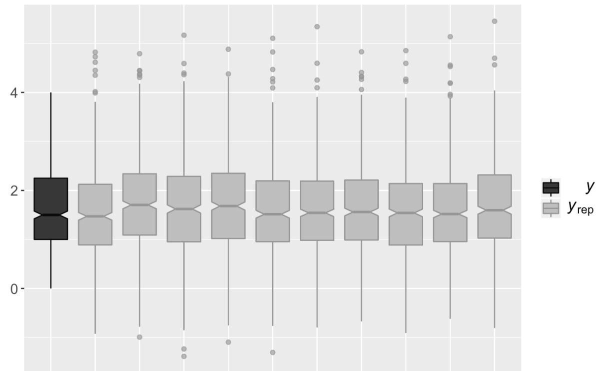


Figure J.22
Kernel Density Plot for Self-Concept in the CCMM Model

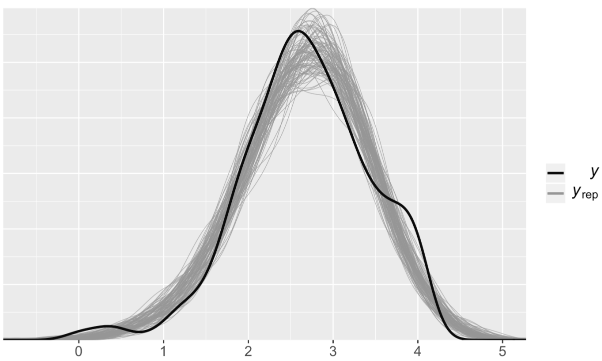


Figure J.23
Boxplot for Self-Concept in the CCMM Model

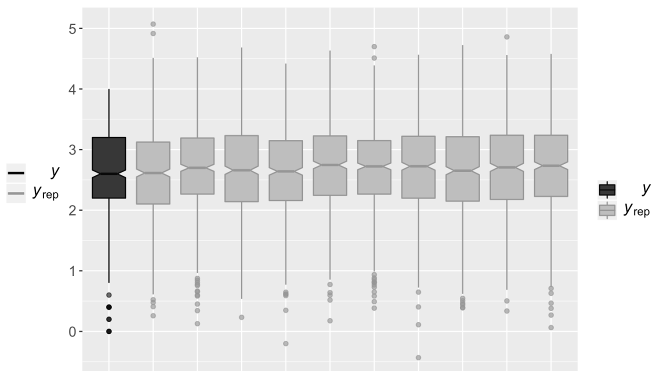


Figure J.24
Kernel Density Plot for Achievement in the CCMM Model

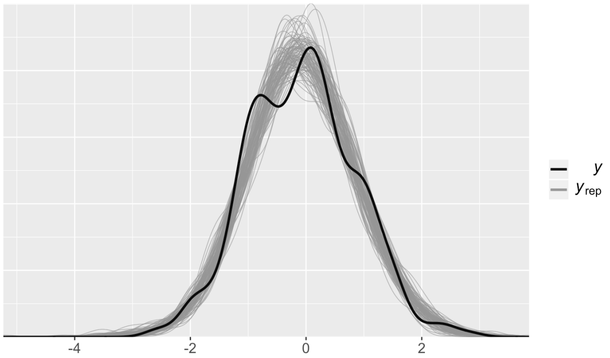


Figure J.25
Boxplot for Achievement in the CCMM Model

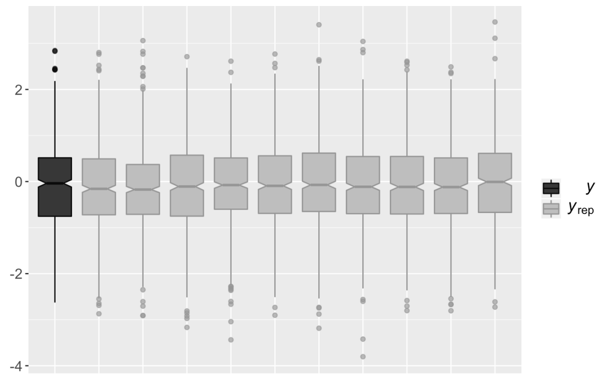
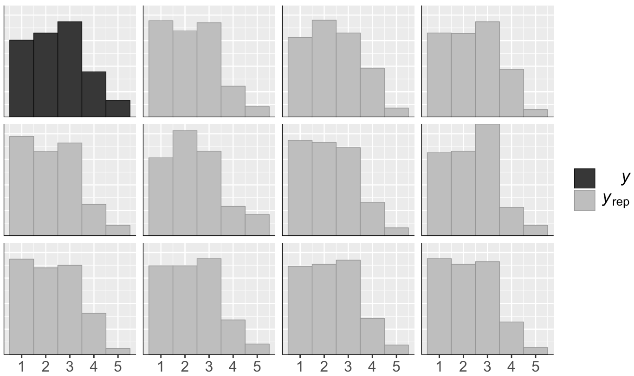


Figure J.26
Histogram for Failure-to-Ability Attribution in the CCMM Model



J.3 The unconditional and conditional CC and CCMM models with brms default priors only

Table J.1

Results of the Two-Level Cross-Classified Random Intercept Model Predicting Students' Beliefs and Achievement Using the Single-Teacher Sample

	Fixed mindset	Mastery goal	P.-approach	P.-avoidance	Attribution	Self-concept	Achievement
Fixed Effects							
Intercept	1.17 [1.05, 1.28]	3.25 [3.17, 3.32]	1.34 [1.22, 1.47]	1.59 [1.48, 1.71]	-	2.70 [2.55, 2.86]	-0.01 [-0.28, 0.26]
Random Effects for Teacher (Number of levels: 17)							
Intercept	0.10 [0.01, 0.25]	0.05 [0.002, 0.14]	0.08 [0.003, 0.22]	0.07 [0.002, 0.20]	0.32 [0.02, 0.66]	0.20 [0.02, 0.40]	0.19 [0.01, 0.53]
Random Effects for Classroom (Number of levels: 24)							
Intercept	0.10 [0.01, 0.24]	0.04 [0.002, 0.12]	0.09 [0.003, 0.24]	0.09 [0.004, 0.25]	0.25 [0.01, 0.55]	0.16 [0.01, 0.35]	0.49 [0.24, 0.78]

Note. N = 342. P.-Approach = Performance-approach goal; P.-avoidance = Performance-avoidance goal; Attribution = Failure-to-ability attribution. Non-informative and weakly informative brms default priors were used. Thresholds for attribution are: 1) -0.95 [-1.23, -0.67], 2) -0.13 [-0.39, 0.13], 3) 0.86 [0.59, 1.15], 4) 1.61 [1.32, 1.93].

Table J.2

Results of the Two-Level Cross-Classified Multiple-Membership Random Intercept Model Predicting Students' Beliefs and Achievement Using the Entire Sample

	Fixed mindset	Mastery goal	P.-approach	P.-avoidance	Attribution	Self-concept	Achievement
Fixed Effects							
Intercept	1.17 [1.08, 1.26]	3.23 [3.19, 3.28]	1.48 [1.37, 1.59]	1.68 [1.58, 1.79]	-	2.69 [2.58, 2.80]	-0.16 [-0.42, 0.10]
Random Effects for Teacher (Number of levels: 29)							
Intercept	0.08 [0.003, 0.19]	0.02 [0.001, 0.05]	0.10 [0.01, 0.25]	0.08 [0.004, 0.23]	0.31 [0.11, 0.49]	0.21 [0.07, 0.33]	0.17 [0.01, 0.51]
Random Effects for Classroom (Number of levels: 34)							
Intercept	0.11 [0.01, 0.21]	0.01 [0.0004, 0.04]	0.17 [0.02, 0.31]	0.15 [0.01, 0.28]	0.13 [0.01, 0.34]	0.09 [0.004, 0.24]	0.70 [0.52, 0.93]

Note. N = 664. P.-approach = Performance-approach goal; P.-avoidance = Performance-avoidance goal; Attribution = Failure-to-ability attribution. Non-informative and weakly informative brms default priors were used. Thresholds for attribution are: 1) -0.79 [-0.97, -0.62], 2) -0.001 [-0.16, 0.16], 3) 0.89 [0.72, 1.07], 4) 1.68 [1.47, 1.89].

Table J.3

Results of the Two-Level Cross-Classified Random Effects Model Predicting Students' Beliefs and Achievement Using the Single-Teacher Sample

	Fixed mindset	**Mastery goal**	**P.-approach**	**P.-avoidance**	**Attribution**	**Self-concept**	**Achievement**
Fixed Effects							
Intercept	1.18 [1.08, 1.29]	3.20 [3.10, 3.29]	1.35 [1.23, 1.48]	1.59 [1.47, 1.72]	-	2.66 [2.57, 2.75]	-0.04 [-0.20, 0.12]
Sex	0.05 [-0.17, 0.27]	**0.19 [0.01, 0.37]**	-0.01 [-0.26, 0.22]	0.11 [-0.11, 0.32]	0.27 [-0.18, 0.74]	-0.08 [-0.28, 0.12]	0.08 [-0.14, 0.28]
Lagged outcome	**0.33 [0.20, 0.46]**	**0.34 [0.22, 0.47]**	**0.45 [0.28, 0.61]**	**0.41 [0.27, 0.56]**	**0.18 [0.08, 0.29]**	**0.64 [0.53, 0.76]**	**0.69 [0.58, 0.82]**
Task	*-0.12 [-0.26, 0.03]*	-0.03 [-0.12, 0.06]	-0.002 [-0.21, 0.19]	0.09 [-0.10, 0.28]	-0.03 [-0.34, 0.29]	0.02 [-0.10, 0.16]	0.01 [-0.15, 0.17]
Evaluation	0.02 [-0.18, 0.23]	-0.002 [-0.11, 0.11]	*0.12 [-0.07, 0.29]*	-0.03 [-0.26, 0.21]	0.05 [-0.33, 0.40]	*0.08 [-0.07, 0.24]*	*-0.09 [-0.25, 0.07]*
Little self-reg.	-0.02 [-0.18, 0.15]	*-0.07 [-0.17, 0.03]*	-0.07 [-0.26, 0.11]	0.02 [-0.17, 0.19]	-0.15 [-0.53, 0.19]	-0.03 [-0.14, 0.08]	0.02 [-0.11, 0.16]
Value	**0.19 [0.03, 0.35]**	-0.002 [-0.13, 0.12]	0.05 [-0.15, 0.25]	0.05 [-0.15, 0.28]	*0.31 [-0.06, 0.63]*	**-0.16 [-0.30, -0.02]**	0.03 [-0.15, 0.20]
Little risk	*0.15 [-0.08, 0.39]*	*-0.12 [-0.26, 0.003]*	**0.22 [0.02, 0.44]**	0.11 [-0.11, 0.34]	0.19 [-0.25, 0.62]	-0.02 [-0.16, 0.12]	*-0.13 [-0.30, 0.05]*
Aggregated FAPs	0.24 [-0.21, 0.67]	-0.32 [-0.73, 0.10]	0.22 [-0.34, 0.78]	-0.02 [-0.59, 0.56]	0.22 [-1.11, 1.48]	0.08 [-0.33, 0.47]	0.18 [-0.48, 0.82]
Random Effects for Teacher (Number of levels: 17)							
Intercept	0.08 [0.003, 0.21]	0.10 [0.01, 0.24]	0.09 [0.004, 0.25]	0.09 [0.004, 0.24]	0.29 [0.01, 0.73]	0.07 [0.003, 0.18]	0.17 [0.01, 0.40]
Sex	0.18 [0.01, 0.48]	0.18 [0.01, 0.43]	0.19 [0.01, 0.55]	0.13 [0.01, 0.38]	0.37 [0.02, 0.98]	0.19 [0.01, 0.47]	0.15 [0.01, 0.45]
Lagged outcome	0.10 [0.01, 0.27]	0.12 [0.01, 0.31]	0.19 [0.01, 0.40]	0.13 [0.01, 0.34]	0.12 [0.01, 0.39]	0.08 [0.004, 0.22]	0.09 [0.01, 0.25]
Task	0.10 [0.004, 0.28]	0.05 [0.002, 0.14]	0.19 [0.01, 0.48]	0.14 [0.01, 0.41]	0.24 [0.01, 0.68]	0.09 [0.004, 0.26]	0.13 [0.01, 0.37]
Evaluation	0.19 [0.01, 0.44]	0.07 [0.003, 0.20]	0.10 [0.004, 0.28]	0.20 [0.01, 0.52]	0.28 [0.01, 0.78]	0.11 [0.01, 0.31]	0.11 [0.01, 0.31]
Little self-reg.	0.13 [0.01, 0.34]	0.08 [0.003, 0.21]	0.16 [0.01, 0.41]	0.12 [0.004, 0.36]	0.33 [0.01, 0.80]	0.07 [0.003, 0.19]	0.10 [0.003, 0.26]
Value	0.10 [0.01, 0.29]	0.11 [0.01, 0.26]	0.12 [0.01, 0.34]	0.13 [0.01, 0.37]	0.28 [0.02, 0.75]	0.09 [0.004, 0.25]	0.17 [0.01, 0.40]
Little risk	0.24 [0.02, 0.56]	0.10 [0.01, 0.25]	0.12 [0.01, 0.35]	0.15 [0.01, 0.44]	0.38 [0.02, 0.97]	0.07 [0.003, 0.21]	0.12 [0.01, 0.33]
Random Effects for Classroom (Number of levels: 24)							
Intercept	0.07 [0.002, 0.20]	0.08 [0.01, 0.19]	0.11 [0.01, 0.27]	0.11 [0.01, 0.28]	0.30 [0.02, 0.70]	0.07 [0.003, 0.19]	0.17 [0.01, 0.37]
Sex	0.21 [0.01, 0.50]	0.15 [0.01, 0.38]	0.16 [0.01, 0.47]	0.13 [0.01, 0.38]	0.39 [0.02, 0.97]	0.15 [0.01, 0.39]	0.14 [0.01, 0.40]
Lagged outcome	0.08 [0.004, 0.24]	0.09 [0.004, 0.23]	0.14 [0.01, 0.36]	0.11 [0.01, 0.27]	0.19 [0.01, 0.48]	0.08 [0.002, 0.21]	0.10 [0.01, 0.25]
Task	0.09 [0.003, 0.25]	0.06 [0.002, 0.17]	0.14 [0.01, 0.39]	0.12 [0.01, 0.34]	0.26 [0.01, 0.67]	0.09 [0.004, 0.25]	0.11 [0.01, 0.30]
Evaluation	0.15 [0.01, 0.39]	0.07 [0.003, 0.20]	0.09 [0.004, 0.27]	0.15 [0.01, 0.42]	0.31 [0.02, 0.78]	0.12 [0.01, 0.31]	0.12 [0.01, 0.32]
Little self-reg.	0.17 [0.01, 0.39]	0.07 [0.003, 0.19]	0.13 [0.01, 0.33]	0.11 [0.004, 0.30]	0.35 [0.02, 0.77]	0.06 [0.003, 0.18]	0.10 [0.004, 0.26]
Value	0.10 [0.003, 0.28]	0.08 [0.003, 0.21]	0.15 [0.01, 0.38]	0.11 [0.01, 0.32]	0.19 [0.01, 0.54]	0.08 [0.004, 0.23]	0.13 [0.01, 0.33]
Little risk	0.21 [0.01, 0.52]	0.09 [0.01, 0.23]	0.11 [0.01, 0.32]	0.14 [0.01, 0.38]	0.40 [0.02, 1.03]	0.09 [0.003, 0.25]	0.10 [0.004, 0.28]

Note. N = 326. P.-approach = Performance-approach goal; P.-avoidance = Performance-avoidance goal; Attribution = Failure-to-ability attribution; Task = Unsophisticated task differentiation; Evaluation = Public evaluation; Little self-reg. = Little promotion of self-regulation; Value = Differential value; Little risk = Little promotion of risk-taking. Non-informative and weakly informative brms default priors were used. Student-level control variables and the group-level predictor were centered using grand-mean centering (CGM). The student-level FAPs dimensions were centered using within cluster means (CWC). Thresholds for attribution are 1) -0.82 [-1.17, -0.48], 2) 0.18 [-0.15, 0.53], 3) 1.42 [1.07, 1.81], 4) 2.41 [2.00, 2.86].

Table J.3

Results of the Two-Level Cross-Classified Random Effects Model Predicting Students' Beliefs and Achievement Using the Single-Teacher Sample

	Fixed mindset	**Mastery goal**	**P.-approach**	**P.-avoidance**	**Attribution**	**Self-concept**	**Achievement**
			Fixed Effects				
Intercept	1.18 [1.08, 1.29]	3.20 [3.10, 3.29]	1.35 [1.23, 1.48]	1.59 [1.47, 1.72]	-	2.66 [2.57, 2.75]	-0.04 [-0.20, 0.12]
Girl	0.05 [-0.17, 0.27]	**0.19 [0.01, 0.37]**	-0.01 [-0.26, 0.22]	0.11 [-0.11, 0.32]	0.27 [-0.18, 0.74]	-0.08 [-0.28, 0.12]	0.08 [-0.14, 0.28]
Lagged outcome	**0.33 [0.20, 0.46]**	**0.34 [0.22, 0.47]**	**0.45 [0.28, 0.61]**	**0.41 [0.27, 0.56]**	**0.18 [0.08, 0.29]**	**0.64 [0.53, 0.76]**	**0.69 [0.58, 0.82]**
Task	*-0.12 [-0.26, 0.03]*	-0.03 [-0.12, 0.06]	-0.002 [-0.21, 0.19]	0.09 [-0.10, 0.28]	-0.03 [-0.34, 0.29]	0.02 [-0.10, 0.16]	0.01 [-0.15, 0.17]
Evaluation	0.02 [-0.18, 0.23]	-0.002 [-0.11, 0.11]	*0.12 [-0.07, 0.29]*	-0.03 [-0.26, 0.21]	0.05 [-0.33, 0.40]	*0.08 [-0.07, 0.24]*	*-0.09 [-0.25, 0.07]*
Little self-reg.	-0.02 [-0.18, 0.15]	-0.07 [-0.17, 0.03]	-0.07 [-0.26, 0.11]	0.02 [-0.17, 0.19]	-0.15 [-0.53, 0.19]	-0.03 [-0.14, 0.08]	0.02 [-0.11, 0.16]
Value	**0.19 [0.03, 0.35]**	-0.002 [-0.13, 0.12]	0.05 [-0.15, 0.25]	0.05 [-0.15, 0.28]	*0.31 [-0.06, 0.63]*	**-0.16 [-0.30, -0.02]**	0.03 [-0.15, 0.20]
Little risk	*0.15 [-0.08, 0.39]*	*-0.12 [-0.26, 0.003]*	**0.22 [0.02, 0.44]**	0.11 [-0.11, 0.34]	0.19 [-0.25, 0.62]	-0.02 [-0.16, 0.12]	*-0.13 [-0.30, 0.05]*
Aggregated FAPs	0.24 [-0.21, 0.67]	-0.32 [-0.73, 0.10]	0.22 [-0.34, 0.78]	-0.02 [-0.59, 0.56]	0.22 [-1.11, 1.48]	0.08 [-0.33, 0.47]	0.18 [-0.48, 0.82]
			Random Effects for Teacher (Number of levels: 17)				
Intercept	0.08 [0.003, 0.21]	0.10 [0.01, 0.24]	0.09 [0.004, 0.25]	0.09 [0.004, 0.24]	0.29 [0.01, 0.73]	0.07 [0.003, 0.18]	0.17 [0.01, 0.40]
Girl	0.18 [0.01, 0.48]	0.18 [0.01, 0.43]	0.19 [0.01, 0.55]	0.13 [0.01, 0.38]	0.37 [0.02, 0.98]	0.19 [0.01, 0.47]	0.15 [0.01, 0.45]
Lagged outcome	0.10 [0.01, 0.27]	0.12 [0.01, 0.31]	0.19 [0.01, 0.40]	0.13 [0.01, 0.34]	0.12 [0.01, 0.39]	0.08 [0.004, 0.22]	0.09 [0.01, 0.25]
Task	0.10 [0.004, 0.28]	0.05 [0.002, 0.14]	0.19 [0.01, 0.48]	0.14 [0.01, 0.41]	0.24 [0.01, 0.68]	0.09 [0.004, 0.26]	0.13 [0.01, 0.37]
Evaluation	0.19 [0.01, 0.44]	0.07 [0.003, 0.20]	0.10 [0.004, 0.28]	0.20 [0.01, 0.52]	0.28 [0.01, 0.78]	0.11 [0.01, 0.31]	0.11 [0.01, 0.31]
Little self-reg.	0.13 [0.01, 0.34]	0.08 [0.003, 0.21]	0.16 [0.01, 0.41]	0.12 [0.004, 0.36]	0.33 [0.01, 0.80]	0.07 [0.003, 0.19]	0.10 [0.003, 0.26]
Value	0.10 [0.01, 0.29]	0.11 [0.01, 0.26]	0.12 [0.01, 0.34]	0.13 [0.01, 0.37]	0.28 [0.02, 0.75]	0.09 [0.004, 0.25]	0.17 [0.01, 0.40]
Little risk	0.24 [0.02, 0.56]	0.10 [0.01, 0.25]	0.12 [0.01, 0.35]	0.15 [0.01, 0.44]	0.38 [0.02, 0.97]	0.07 [0.003, 0.21]	0.12 [0.01, 0.33]
			Random Effects for Classroom (Number of levels: 24)				
Intercept	0.07 [0.002, 0.20]	0.08 [0.01, 0.19]	0.11 [0.01, 0.27]	0.11 [0.01, 0.28]	0.30 [0.02, 0.70]	0.07 [0.003, 0.19]	0.17 [0.01, 0.37]
Girl	0.21 [0.01, 0.50]	0.15 [0.01, 0.38]	0.16 [0.01, 0.47]	0.13 [0.01, 0.38]	0.39 [0.02, 0.97]	0.15 [0.01, 0.39]	0.14 [0.01, 0.40]
Lagged outcome	0.08 [0.004, 0.24]	0.09 [0.004, 0.23]	0.14 [0.01, 0.36]	0.11 [0.01, 0.27]	0.19 [0.01, 0.48]	0.08 [0.002, 0.21]	0.10 [0.01, 0.25]
Task	0.09 [0.003, 0.25]	0.06 [0.002, 0.17]	0.14 [0.01, 0.39]	0.12 [0.01, 0.34]	0.26 [0.01, 0.67]	0.09 [0.004, 0.25]	0.11 [0.01, 0.30]
Evaluation	0.15 [0.01, 0.39]	0.07 [0.003, 0.20]	0.09 [0.004, 0.27]	0.15 [0.01, 0.42]	0.31 [0.02, 0.78]	0.12 [0.01, 0.31]	0.12 [0.01, 0.32]
Little self-reg.	0.17 [0.01, 0.39]	0.07 [0.003, 0.19]	0.13 [0.01, 0.33]	0.11 [0.004, 0.30]	0.35 [0.02, 0.77]	0.06 [0.003, 0.18]	0.10 [0.004, 0.26]
Value	0.10 [0.003, 0.28]	0.08 [0.003, 0.21]	0.15 [0.01, 0.38]	0.11 [0.01, 0.32]	0.19 [0.01, 0.54]	0.08 [0.004, 0.23]	0.13 [0.01, 0.33]
Little risk	0.21 [0.01, 0.52]	0.09 [0.01, 0.23]	0.11 [0.01, 0.32]	0.14 [0.01, 0.38]	0.40 [0.02, 1.03]	0.09 [0.003, 0.25]	0.10 [0.004, 0.28]

Note. N = 326. P.-approach = Performance-approach goal; P.-avoidance = Performance-avoidance goal; Attribution = Failure-to-ability attribution; Task = Unsophisticated task differentiation; Evaluation = Public evaluation; Little self-reg. = Little promotion of self-regulation; Value = Differential value; Little risk = Little promotion of risk-taking. Non-informative and weakly informative brms default priors were used. Student-level control variables and the group-level predictor were centred using grand-mean centring (CGM). The student-level FAPs dimensions were centred using within cluster means (CWC). Thresholds for attribution are 1) -0.82 [-1.17, -0.48], 2) 0.18 [-0.15, 0.53], 3) 1.42 [1.07, 1.81], 4) 2.41 [2.00, 2.86].

Table J.4

Results of the Two-Level Cross-Classified Multiple Membership Random Intercept Model Predicting Students' Beliefs and Achievement Using the Entire Sample

	Fixed mindset	**Mastery goal**	**P.-approach**	**P.-avoidance**	**Attribution**	**Self-concept**	**Achievement**
			Fixed Effects				
Intercept	1.17 [1.10, 1.24]	3.18 [3.13, 3.24]	1.43 [1.35, 1.53]	1.65 [1.55, 1.75]	-	2.68 [2.62, 2.75]	-0.12 [-0.28, 0.03]
Girl	0.05 [-0.07, 0.18]	0.07 [-0.02, 0.17]	0.05 [-0.14, 0.26]	0.06 [-0.12, 0.24]	0.26 [-0.03, 0.52]	-0.09 [-0.22, 0.04]	0.07 [-0.07, 0.20]
Lagged outcome	**0.37 [0.27, 0.47]**	**0.36 [0.27, 0.45]**	**0.45 [0.35, 0.55]**	**0.44 [0.35, 0.53]**	**0.16 [0.09, 0.23]**	**0.61 [0.52, 0.70]**	**0.67 [0.55, 0.78]**
Aggregated FAPs	0.05 [-0.23, 0.32]	-0.01 [-0.19, 0.18]	**0.52 [0.17, 0.89]**	0.28 [-0.09, 0.66]	0.42 [-0.15, 0.97]	0.01 [-0.26, 0.28]	0.12 [-0.42, 0.71]
			Random Effects for Teacher (Number of levels: 29)				
Intercept	0.06 [0.003, 0.16]	0.05 [0.001, 0.13]	0.08 [0.003, 0.21]	0.09 [0.01, 0.24]	0.16 [0.01, 0.38]	0.06 [0.003, 0.17]	0.11 [0.004, 0.30]
Girl	0.09 [0.004, 0.26]	0.11 [0.01, 0.26]	0.25 [0.02, 0.56]	0.17 [0.01, 0.44]	0.33 [0.02, 0.76]	0.13 [0.01, 0.33]	0.11 [0.01, 0.30]
Lagged outcome	0.11 [0.004, 0.28]	0.08 [0.004, 0.20]	0.11 [0.01, 0.28]	0.10 [0.01, 0.24]	0.10 [0.004, 0.28]	0.08 [0.003, 0.21]	0.15 [0.01, 0.32]
			Random Effects for Classroom (Number of levels: 34)				
Intercept	0.07 [0.003, 0.16]	0.04 [0.002, 0.10]	0.08 [0.004, 0.20]	0.10 [0.004, 0.23]	0.12 [0.01, 0.30]	0.07 [0.004, 0.16]	0.36 [0.23, 0.52]
Girl	0.09 [0.004, 0.26]	0.07 [0.003, 0.20]	0.17 [0.01, 0.42]	0.17 [0.01, 0.41]	0.23 [0.01, 0.59]	0.10 [0.004, 0.26]	0.12 [0.01, 0.31]
Lagged outcome	0.13 [0.01, 0.28]	0.10 [0.01, 0.19]	0.14 [0.01, 0.29]	0.08 [0.003, 0.21]	0.12 [0.01, 0.29]	0.11 [0.01, 0.24]	0.15 [0.01, 0.30]

Note. N = 625. P.-approach = Performance-approach goal; P.-avoidance = Performance-avoidance goal; Attribution = Failure-to-ability attribution. Non-informative and weakly informative brms default priors were used. All predictor variables were centred using grand-mean centring (CGM). Thresholds for attribution are 1) -0.48 [-0.67, -0.28], 2) 0.33 [0.15, 0.53], 3) 1.28 [1.08, 1.5o], 4) 2.14 [1.89, 2.39].

Appendix K

Example Summary of Student Interview Resulting from Data Familiarisation

B=Boy
T=Teacher

- Bs report T changed some behaviours. She did some "wacky" things, like having them all stand up in the lessons. But nothing really changed about teaching per se.
- All students get the same tasks. Bs say that getting easy questions all the time would be pointless because it would not be a challenge. If T would give them easier tasks, they think T would think they are stupid and not good mathematicians → task differentiation affects self-concept
- Bs think they get the same tasks because that makes it easier for T to judge their performance and see what needs to be improved. If all would get different tasks, then performance across students couldn't be compared. It also sounds as though one B says that if she does not give the more difficult test to someone, she does not think they are capable of it → tasks convey teacher expectations
- B says they do a lot of self-assessment, but T never checks their books. Bs say some cheat because of this. They think they should do self-assessment to look at their mistakes and learn from them. Not to make mistakes again and be like teachers → promotion of self-regulation affects mastery goals?
- They also self-evaluate almost every class using a glum face, straight face, happy face, and sensible comment. This is school policy but not all Ts do it. Bs are not sure why they are supposed to do this. They say T pretty much never checks. And one says she can then teach again what they found difficult, but another B says she always says she would not teach anything again.
- Bs get warnings and demerits for chatting or turning around.
- B thinks self-evaluation might go into their assessment as well. And students learn themselves where they are at. But T does not seem to have communicated why she does it.
- T asked Ss at the beginning of the year to set their own learning goals. But then she did not seem to engage with that. Now T sets their goals. Bs are not overly happy about her setting goals, it seems that she sets very high ones because they say it might be better if she sets some easier ones at the beginning so that all have a chance. → high expectations
- Bs think they don't get much choice. They think it's because she wants them to succeed and work more so that they progress. They think she is thinking about their future and does the job not just for the money → Bs think less choice is related to more progress?
- T doesn't like getting asked the same question repeatedly.
- Mixed opinions regarding ability grouping. Some say it makes sense to have similarly performing students together so that they have a similar understanding and other says mixed is better so that they can help each other. They make these suggestions but are not sure what T actually did. → heterogenous pairs associated with mastery?
- Bs also discuss where weak students should sit. They think it's good if they all sit together, best at the front, because that's the most important place. One B explains

how there were different groups in primary school, clearly marking ability, and that that would be good. So it seems it's what they are used to.

- Sometimes, usually before holidays, class does something fun, some group work, and students can choose with whom to work.
- Bs describe that T usually picks those that are a bit weaker or don't know the answers. Only if nobody knows and many tried guessing, she has specific students she goes to because they are likely to get it right. I think this nicely mirrors what T told me. She says it kind of trickles down; when the questions get very tricky there are only a few students left that participate. Bs say there were days were only the strong ones could do it and that makes them feel angry, unhappy and underpowered. Bs explicitly say T should try to also pick those that don't have their hand up because those might know it but they are just a bit shy.
- Bs say that if you give an answer and it is correct, it will stick with you. So you should not take the people who find it easy anyways. → differential value affects mastery?
- T goes around to check if all are involved in pair work. If she thinks not, she asks them if they are involved → T encourages all to participate.
- Bs say that when someone makes a mistake, T just says "Noooo" and seems to wait for a while or move on to next student. Bs say that T does not discuss mistakes. This makes them feel dumb, not as intelligent and as though they have not learned anything. They say if you don't know what you did wrong, you cannot learn from it. They also say that if she just says "Noooo" and moves on, she seems disappointed in them, which is why they feel less intelligent. They say she says if something is wrong but does not do it in a mean way. But the absence of discussion (and probably normalising the mistake) seems enough to make them feel bad. Not all agree though. Some say she doesn't seem disappointed, she just says it's wrong. Her reactions make some feel down, sad, and some say neither good nor bad. → Handling of mistakes affects self-concept.
- Bs say she has very high expectations, expects them to learn a lot in a short period.
- Bs think T does not really care about mistakes.
- T often plays a game where the class has to be faster than her otherwise they get work in text books, which they don't like. Bs say that she is faster but she always lets them win. This is also what T said. She'll always find a way for them to redeem themselves. So not sure if this can really count as criticism for mistakes.
- Speed emphasis. They think it alerts them more, pushes their brain, helps in the lessons because they won't dawdle. → Speed emphasis associated with learning?
- Bs talk about a situation in which they had some kind of competition with speed as well. All Ss had to stand up and if they had the answer lift their hand and T came to check. Then they could sit down if they were correct. All the weak ones were left and felt quite humiliated because they needed more time or couldn't solve it → speed emphasis potentially affects self-concept?
- Bs think T has a growth mindset but cannot explain why.
- Bs say that T does not provide much support. You have to have your hand up and then she gives you an answer but she does not really help, she does not provide more tasks to work on or practice.
- One B said she shows she has a growth mindset because when they make a mistake and then correct it, she praises them for it. So this can explain why reactions to mistakes are related to mindset. Similarly, she asks Ss how to improve their work (I guess rather than just correcting it herself). → Handling of mistake helps to infer T mindset.

Appendix L

Steps in the Coding Process and Prominent Changes in Codes

Table L.1

Steps in the Coding Process and Prominent Changes

Steps in coding process	**Prominent changes in codes**	**Notes and clarifications**
1) First round of coding • Coding of all interviews using the initial coding scheme • Creating new codes based on data	New codes were created: *Random grouping, weak at front, pairing for behaviour, no group or no pair work, demerit, merit, independent learners, structure, mastery emphasis, challenge students, show workings, foster resilience, performance emphasis, assessment, diagnose performance, speed emphasis, differentiate by time and support, teacher knows best, teacher-student relationship, teacher cares, discipline, increase focus/concentration, school policy, observation effect, situation attribution, learning at a different pace, student emotion (annoyance, boredom, disappointment, don't care, embarrassment, fairness, happy, laziness, pressure, pride, rejections, sadness, worry, humiliation), nature of maths*	No instances detected of: • *multiple products/solutions* • *one product/solution* • *equal attention/interest* • *different attention/interest* • *fixed mindset* • *person-oriented recognition*
2) Refinement of codes • Merging or re-naming codes where appropriate • Excluding codes if there is no clear relationship to the study or if they are only very few instances	• Differentiated *different tasks* from *staggered tasks student choice* • Merged *same amount, same difficulty* and *same task* into *same task* • Split *process/person/product recognition* into *praise* and *criticism* • Collapsed *demerit* and *merit* under *praise* and *criticism* • Deleted *competitive/non-competitive rewards* as not prevalent • Changed *diagnose performance* to *formative assessment* • *Assessment* changed to *formal assessment* • Split *help-seeking encouraged* and *discouraged* into *teacher* and *peer support encouraged* and *discouraged* • Merged *call another student, moving on, no discussion of mistakes* and *T ignores mistakes* into *moving on*	• Note that *staggered tasks student choice* is of course related to *choice* but these instances were only coded as *staggered tasks student choice*

	• Merged *waiting, trying again* and *second chance* into *try again/self-correction* • Collapsed all *positive emotions* into one code • Collapsed all *negative emotions* into one code • Merged *learning at a different pace* with *ability attribution* • Deleted *discipline* and *maths enjoyment* because it is not related to the mindset framework and only few instances were identified • Deleted tea*cher cares, student teacher relationship,* and *teacher knows best* because there were very few instances	
3) Second full round of coding • Coding all interviews using the adapted coding scheme • Making additional changes to the scheme where appropriate, for instance re-naming codes • Developing descriptions and selecting examples to create complete coding scheme	• Found instances of *fixed groups* • Coded an instance of *different product* • Renamed *don't care* to *indifferent* • Renamed *most participate* to *encouraging all to participate* as the teacher might try but not be successful • Split *self-evaluation* (only encompassing reflections on learning and progress) from *self- or peer-assessment* (everything related to marking) • Combined *independent learners* with *promotion of self-regulation* • Separated *public evaluation* and *social comparison* into two codes as they are distinct despite some overlap • Removed *no self-evaluation* and instead coded for *self-evaluation* only • Collapsed instances of *set own goals* under *self-evaluation* as there were few instances • Collapsed *equal attention* and *interest* under *encouraging all to participate*	• Sometimes difficult to distinguish *peer support encouraged* and *group and pair work* (*group and pair work* is specifically for group activities; *peer support* is if one student needs help they can ask their partner) • Sometimes difficult to distinguish *growth* from *mastery* (coded anything related to being able to improve and making progress as growth; if the focus is on wanting to improve and grow, it's mastery goal) • *Differentiate by support* is always also *teacher-support encouraged* but not the other way around • Sometimes difficult to differentiate *product-* from *process-oriented praise/criticism* (coded anything related to changes in performance (improve, decline) as process; anything related to amount of work done with time constraint (e.g., work in class) as *product-oriented* but not getting enough done with no time constraint (e.g., work at home) was coded as *process-oriented*)

4)	Coding by a second coder • Approximately 10% of the data are coded by a second coder • Reliability analysis	• Second coder detected some instances I have overlooked and vice versa • Second coder misunderstood *performance goals, self-concept*, and *public evaluation*	• More information is provided in the Section 8.2.1.1
5)	Adjustment of coding as required based on step 4	• Coded some instances I previously overlooked (e.g., *mostly strong participate, formative assessment*) and checked all transcripts for these codes.	• Overall only few adjustments made

Appendix M

Coding Framework for Fixed-Ability Practices and Motivational Beliefs

Note that codes that cannot be classified as neither non-ability based nor ability-based are presented across the middle of the two columns. S = student; T = teacher.

Table M.1

Coding Framework for Fixed-Ability Practices

Category	Fixed-ability codes	Description and example	Non-fixed-ability codes	Description and example
Tasks				
Unsophisticated task differentiation	Different tasks	-T assigns different tasks based on ability/performance -if tasks are staggered by difficulty level, Ss cannot choose what tasks to do "…depending on what mark you do, you do a bit of, er, like, three questions on whatever mark you got."	Same task	-T gives the same tasks/work sheets to all Ss "She gives us the same task."
			Staggered task student choice	-T hands out the same worksheets to all Ss but Ss can choose which difficulty level to work on "…if you wanna challenge yourself go to the, straight to the high ones as well."
Unidimensional tasks	One method	-T wants Ss to use only one specific method to solve a problem "So we should be passing that on and not just saying 'do it how you wanna do it.'	Different methods	-T shows or allows different ways of solving a problem "And to show them, that that is actually the same process, look, or the same result gotten by a different process…"

		Because if there are better and worse ways of doing it…”		
			Different solutions/ products	-Tasks that can lead to different solutions or Ss can calculate with different numbers “Like, you can make it quite, you can do that really simply 1 times 2/5 or you can think about a really complex fraction...”
			Extension task	-T provides tasks for Ss who finished early -extra challenge in addition to the normal class content “…it shouldn’t be something that if they can’t do the star task, they can’t meet the objective.”
Grouping				
Group and pair work	No group or pair activities	-there is no group work OR -there is no pair work “We don't, we don't really do group work at all.”	Group or pair activities	-Ss engage in group or pair work -activities explicitly meant for group or pair work “I love the mini white boards. Well it's nice because they can talk.”
Seating and ability grouping		Random seating plan	-there is no particular reason that certain Ss are grouped or seated together, such as alphabetic order “Pretty much alphabetic and, um, boy-girl-boy-girl.”	
		Seating for behaviour management	-Ss are sat together for behaviour management “It's not for us to talk with our friends because she puts as with people that she knows we're not gonna talk to.”	
		Partner/group choice	-Ss can decide with whom to work or whom to sit next to “And those groups they can usually choose.”	
	Homogenous pairs	-Ss at the same ability level work/sit together “Happy because like say me and [Student name] have the same score, if I'm with him,	Heterogeneous pairs	-Ss who have a different performance levels work/sit together

		then I think we'd be able to be able to talk about how we are feeling about it."		"And in some instances I will have a stronger and weaker one…"
	Weak at front	-Ss that are struggling are seated at the front of the class "So I've moved her at the front so I can give her a bit more attention so I remember to go and check on her."		
	Ability setting	-all references made to ability sets such as Ss belonging to different sets or differences between sets "Like, at the beginning of the year, they put us in slightly random sets 'cause I wasn't too good when I was in B…"		
Stability	Fixed pairs/sets	-Ss cannot change ability sets (school wide) or the pairs/groups/seating within classes "…if you are in a different school and you're in the lowest set for the whole year and you know, you're, like, you're supposed to be in the top set…"	Fluid pairs/sets	-Ss can change ability sets (school wide) or the pairs/groups/seating within classes "Because we change sets every half term we also have new spaces."
Evaluation				
Public evaluation	Public evaluation	-T discusses Ss' performance or behaviour in front of the class -public praise "… sometimes people don't want to be, like, have the class knowing that they are confused…"	Private evaluation	-T gives praise in private, such as quietly to S or directly in their maths books -T enters merits at the end of class without Ss present "Demerits, she comes to you personally."
	Social comparison	- T compares Ss' performance -anything related to competition between Ss or classes -Ss comparing each other "…sometimes um she'd compare one student to another."		
Performance over process emphasis	Performance emphasis	-T is mostly concerned with performance -correct answers are more important than effort	Mastery emphasis	-T focuses on deep rather than surface understanding

	“In my opinion I think that she thinks more about the correct answers not the effort going into it.”		-understanding concepts rather than following procedures “So, to develop the mastery rather than just the fluency […]
Speed emphasis	-T emphasises, values or rewards speed “And if you can do it relatively quickly then you're gonna enjoy and the lesson makes you feel clever kind of thing.”	Effort emphasis	-T emphasises and values the importance of effort “She'll like, kind of congratulate you for trying your hardest.”
		Challenge students	-T challenges Ss and stretches their understanding “We've got to challenge them but it, it's part of maths.”
Formal assessment	-everything relating to formal tests and assessment “Yeah there is a lot of pressure to keep on doing tests.”	Formative assessment	-T monitors learning to provide feedback but also to adjust their teaching -T checks Ss’ understanding during normal lesson “…if you see the majority of people in the class did the easiest you know that we're all finding it hard so you can do another lesson on that.”
		Show workings	-T wants Ss to explain their reasoning, their steps in solving a problem -T wants that Ss write out their working properly “…I'm not interested in the answer. I'm more interested, I'm interested in the working out.”
		Foster resilience	-T wants to foster resilience, that is their ability to persist and bounce back after failure “But the ability to be resilient, not giving up, um when it's getting difficult…”
		Process-oriented praise	-recognition/praise/merits for effort, working hard, strategy, doing extra work etc. -praising improvement and progress

				"That's because you have, like, she likes to praise the improvements…"
			Process-oriented criticism	-demerits/criticism or punishment for process and misbehaviour, such as not listening or not completing homework (without time constraints) -criticism for not trying hard or declining performance "If they are not listening it will be a warning…"
	Product-oriented praise	-recognition/praise/merits for performance, outcomes, correct solution, good work, the amount of work done "Merits, for example she gave me an, a merit for getting all of my MAD time correct."		
	Product-oriented criticism	-demerits/criticism or punishment for products such as getting wrong answers -not having enough work done (with time constraints, e.g., class work as opposed to homework) "And there are kids that, like, focus on it, like, revise it every day and then still get in trouble…"		
Little autonomy				
Little choice	No student choice	-Ss have to do everything T's way -no options "Everything's basically Miss's way."	Students have choice	-Ss can work in their own way and make decisions -Ss can choose which tasks to work on (excludes staggered tasks) "Well we learn it but then we are allowed to do, like, our own way sometimes."
	Structure	-everything related to how structured the classes are (e.g., classes are highly structured		

		with pre-set times for tasks and the order of doing activities) "…I find routines nice 'cause then you don't have to be questioning what's gonna happen next."		
Little promotion of self-regulation	Teacher sets goals	-T sets goals and targets "…she tells us what our targets have to to be by the end of the half term"	Self-evaluation	-Ss are encouraged to reflect on their learning, their strengths and weaknesses -Ss evaluate their own work (excludes checking for correctness) and set their own goals "…if we know what went well we could like look back from that and say well I can be better now."
			Self- or peer-assessment	-Ss correct and mark their own and/or each other's work "But when they mark each other, so they will have to look at the question…"
Differential value	Mostly strong participate	-usually the stronger Ss are called on more often "She just goes straight away to the person who knows it."	Encouraging all to participate	-All Ss are encouraged to participate in the class -people are called on equally -calling on those that were inattentive "…she wants people to, if they're unsure, like, have a go at answering questions."
Little promotion of risk-taking				
Little encouragement of help-seeking	Teacher support discouraged	-T does not help Ss properly when they have questions -Ss don't feel safe to ask for help "… half the time she gets, like, quite angry because, um, she thinks that we are attention seeking…"	Teacher support encouraged	-T helps Ss -T takes time to answer questions -Ss feel safe to ask T for help "But now I'm, like, actually OK to ask her because she's made me feel OK to ask her things."
	Peer support discouraged	-T does not encourage Ss to help each other -Ss have to ask for permission if they want to help others "We only have to do the work by ourselves."	Peer support encouraged	-T allows Ss to help each other when working on tasks or Ss refer to helping each other - T promotes cooperation among Ss "And they can help each other.."

			Differentiate by support or time	-T gives more support or time to weaker Ss “We all get the same tasks except she helps a little more, like, people a bit more.”
Unconstructive handling of mistakes			Mistakes for learning	-direct messages about the benefit of mistakes -using mistakes for learning “…she would say, like 'I'm so happy that you made that mistake because then everyone else can learn from it.'”
	Critical of mistakes	-T shows disappointment or anger -T humiliates or laughs -T punishes S for mistakes (e.g., demerits) “It's like you are afraid to talk because if you do a mistake you get a demerit.”	Mistakes are OK	-absence of negative T reactions -inhibition of negative S reactions -Ss think T does not mind mistakes “She would say 'not quite' or if you, if you are nearly there she'll say 'nearly'.”
	Moving on	-T swiftly moves to another S when a mistake occurs -T points out mistakes but does not engage in discussion -T ignores the mistake “She'd just say 'No' and goes to the next person.”	Discuss mistakes	-T discusses mistakes such as showing the working and explaining where S went wrong “And if we, like, get it wrong she explains it, like, what we have done wrong.”
	T corrects mistakes	-T gives the correct solution “… she would just tell the answer 'cause they have to move on to the next part.”	Try again/self-correction	-T gives Ss second chances and time to try again “And if we do get it wrong she tells us to look back at it again.”
Contextual codes		School policy	-anything relating to what is done at the school rather than what the T decided to do (background knowledge needed for coding) “Um, because they have to go up or maintain their score each week.”	
		Observation effect	-anything pertaining to observations affecting T and Ss “She acted a little bit more, like enthusiastic about, yah, professional about the job.”	

Table M.2

Coding Framework for Motivational Beliefs

Category	More growth code	Description and example	More fixed code	Description and example
Mindset				
Growth and fixed mindset	Growth	-messages/belief about being able to improve -or anything related to having improved and gotten better, having made progress "Because she wants you to improve on your intelligent in maths."	Fixed	-messages/belief about ability being fixed or stable with little hope of improvement "And the fact that someone is on a Level 3 and someone else is on a 5, well that's just what it is."
Achievement goal orientation				
Mastery goal, performance approach goal, and performance-avoidance goal	Mastery goal	-S focuses on learning, understanding and mastering the task -S wants to develop or improve skills and competences -S wants challenges "I wanna do harder work 'cause then you are not really learning much at all."	Performance-approach goal	-Ss focus on looking competent and smart -outperforming others "Like, personally I love it. Because then everyone knows I got it right!"
			Showing off	-Ss brag about their performance "…it gives you, the people that got it right, like a chance, like, brag about it…"
			Performance-avoidance goal	-Ss focus on avoiding looking dumb in front of others -negative emotions if Ss perform poorly in front of others "And everybody put their hands up and I was just like I'm, I'm gonna put my hand up, 'cause like, I feel, I feel embarrassed"
			Cheating	-anything related to cheating "…if you put it on the back like you get some people that turn it over and copy all the answers."

Attribution				
Effort attribution, ability attribution, and situation attribution	Effort attribution	-performance is explained by effort -effort can lead to improvement “Yah, their grades are usually, I mean the effort is usually worth the result.”	Ability attribution	-performance is explained by ability “Maybe she knows that you are a bit more cleverer and that, er, she knew you would be able to push yourself to that.”
		Situation Attribution	-performance is explained by the situation or context “If she couldn't get it, and then which means there is something I haven’t explained it.”	
Students’ self-concept				
	Positive self-concept	-Ss think they are good at maths -Ss have high expectations for themselves “’Cause that means I know that I'm going well.”	Negative self-concept	-Ss think they are bad at maths -Ss have low expectations for themselves “Makes me feel not as intelligent.”
Teachers’ perceived ability and expectations for students				
	High expectations	-T has high expectations for all Ss “I think I have a reputation for stretching them till they just about snap.”	Low expectations	-T has or seems to have low expectations for Ss “And Set 4 is unlikely they're going to choose A-level maths”
			Differential expectations	-different expectations for different Ss or sets “It's just because she's, like, teaching a higher, like the highest set of the year, she, like, she kind of has higher expectations.”
Nature of maths				
	Complex nature of maths	-comments about what maths is about -e.g., maths is about thinking and cognitive activity, about understanding and not just memorising and following a recipe -maths is about struggling “They haven't thought, they've done a recipe and maths isn't a recipe.”		

Student emotion		
	Positive emotion	-any positive emotions, including pride, happiness, enjoyment, feeling good etc. "I love where Miss asks that!"
	Negative emotion	-any negative emotions including disappointment, embarrassment, sadness, rejection, annoyance, worry, pressure etc "It's kind of depressing."
	Indifferent	-S does not really care about what T does, feels indifferent, or thinks it's alright "I don't care about merits."
	Fairness	-any comments about whether something is fair or unfair, such as a T action "It's quite unfair."

Appendix N

Examples for a Teacher and Student Matrix

Table N.1

Example of Teacher Matrix Linking Teachers' Motivational Beliefs and Other Justifications to Fixed-Ability Practices

	Extracts			
Fixed-ability practices	**Ms Evans (more fixed)**	**Ms Walker (more fixed)**	**Ms Barns (less fixed)**	**Ms Green (less fixed)**
Public evaluation	T: On that [points on white board behind her] that's like the ranking for the year [inaudible]. And that's eff- [effort], er that's by score for, like, general, it's not, it's not specifically effort. It's like I: Yeah. T: =effort, participation, politeness, I: I see. T: =behaviour, equipment. Everything that you would want that's not attainment. --- I: OK. Alright, aaah [both laugh]. OK, so again, get on with this. OK, so there were a couple of situations where you asked, er, the students to kind of put their hand up if they got all the T: Yeah! I: =answers correct. T: Yeah. ---	I: Um, I've noticed that you often point out when pupils made progress. So you would say '[Girl name], well done, that's big progress for you' or something. Why, what's your intention behind that? Why are you telling them? T: Um, what was obviously, I wanted to be clear that I'm rewarding progress not just you're getting all the answers right. Um I: Why? T: Um, because they need to realise that everyone is a slightly different ability. And it's, you will get there if you put in the hard work. And sometimes I think the students that have found it harder and who have got there, deserve to be praised more than the ones that	I: OK. And do you let them know as well? That you// T: Well at the time I'm saying what I'm doing. I: OK. Alright. Mm-hm! T: Um, and I think I'm a lot more open about the demeriting than the meriting in actual fact. I: Oh you mean you would, OK. So demerits you would tell them more often than if you'd give them a merit [they might not even know it.] T: [Yah because sometimes I think] I havn't - Oh that was a good lesson and I'll sit down and think well now, who tried hard? I: Mm. [And you have to do that] T: [And I go click click click] I: You have to do that every lesson as well? T: No, I don't do it every lesson just now and then. ---	T: And as long as I didn't notice you and I I I think you are doing um some good work [I: Ah ok.]. So this standards I I would say it's very low [I: Mh hm]. If I'm strict there will be lots of children who won't be able to get house merits and um...but...on the other hand I want to encourage more people [I: Mh.] to do well. And so what I think um... the house points for them, valuable to them, is they are having um they are having houses, houses they have different, 5 different houses [I: Ok] and then they have house assembly. And they gonna compare which house has more points. Um every week and they're competing. I: Every week? Wow [laughs]. T: Yeah! So they know it's it's important but house points they do not count the demerits. They only count the positive merits. --- I: So, when pupils are working on the task. So they are doing kind of

T: Er, it's just a general, like, a really quick snap shot of, like, am I going too fast, am I going too slow. I: Mh-hm. T: Have they all nailed this. Just a couple of times a lesson just to literately check. I, they, they could all be lying but they are generally not and, um, you'd, if I, and, like, it's also useful if they are about to be set off for a 15 minute independent task, to quick clock, OK, she didn't put her hand up, she didn't put her hand up, go to them first. Check that they are OK.

I: Cool. Um, and then...Oh yeah after the knowledge quiz on Friday T: Mh-hm. Yah. I: =you asked the students to kind of read out the score that they've just got on the T: Quiz, yah. I: =on the quiz. T: Yah. I: Why, why did you do that? T: So, that's a department policy. I: OK. T: Um, because they have to go up or maintain their score each week. So they do that quiz every single week I: Mh-hm. T: =um, and it's an accountability measure to make sure that they are have found it really easy straight away. It doesn't mean the others shouldn't be praised but that's obviously a really good skill that they've developed. Like resilience, which is important. Um, I also, in terms of, like, the class dynamic and which kids are considered to be cooler and which kids get followed more and which kids the class I know considers to be like dumb almost and the ones which the class considers to be like the clever ones. I like to try, like, hack into that dynamic by doing things like that. So for example [Girl name], um, will often, she is really good at asking, um, questions, which is great, but she'll often ask questions which are wrong. In which case the class then finds it kind of funny and things like, so it's a chance for me to be, like, 'don't be thinking [Girl name] is not really clever because she is'. Um, and then that's the same with [Girl name]. I say that to her quite a lot.

T: So there I was writing their names [to give the merits.] I: [Exactly, so this is] I: Ah OK. That's what it's f-, you, OK. So you write the name down for merits or wha-, what T: Yah. I: =that was my question. Why are you writing it? T: Why am I writing it, because I won't remember. I: Ah OK. Good. So bascially they...everyone who has finished [will get] T: [everyone would...yah] I: Yah. Why? [I laughs] [...] [School bells ring] T: Why? Because they had managed// I: Oh already? Oh! T: It's alright, I don't teach this period. I: Oh OK! [laughs] Sorry [whispered]. [both laugh] T: I know. Um, because I've got to have some way of rewarding. Um... I: OK. And just want to recognise them for... T: Yah. Yah. I: OK. T: Because they, er, yeah can't have stopped to chat along the way. Other kids...probably have. Possibly have. I: So... most likely or possibly the ones who have finished were the ones [who have, like, worked most concentrated.] T: [who really engaged, yah]

individual work or they are speaking to your partners. I've noticed that a couple of times you would, like you are going through the class and then you're kind of commenting on how, like the pupils are doing. So sometimes you would say something like um 'Oh like um, some, some students have already finished four tasks and others are still working on two tasks' or something. T: Yeah.

T: Um...because they don't see each others' progress. They don't have comparison and some people tend to be lazy. And also when I say it, they notice the teacher is observing them. I'm watching them. So it's kind of um telling them 'you better work harder 'cause I'm watching you!' Lots of children are, they need ah, they need teachers... kind of push from time to time. Otherwise they would do as little [I: Mhm] as possible. Ah it's very difficult for me to do this all the time [I: Mhm]. So I can only do it, so when I can. That time is sometime I don't have to answer their question, so I have more time to tell 'you havn't done enough' [T laughs and I joins]. But when I started to help them in more details individually, it would be hard for me to...[I: Yeah.] to talk about it at that time.

I: Alright, so basically just to kind of show them where they are standing and

	actually putting in the effort to improve. --- T: That takes so much more time. Yeah, that's only because it takes a lot of time. I: OK. Do you think that this could create a sense of competition though between the students? T: Yeah but they also get, they, they know, they know it's a competitive environment because they know that they are set, that they are re-set every half term based on how they do. I: Mhm. T: Um, and I think two, like the two things, there is like culture of error and, like, that it's fine to get things wrong. There's nothing wrong with not getting things right because we all, it happens all the time. We all make mistakes. I make mistakes. Tell me if I've made a, I made a mistake today. I was like 'Oooh! Oh I see what I've done. Yah, good.' That's a good, because it's always good for learning if you make mistakes because that's a learning point. Um, so it's nothing wrong with not being, not	Because she came up from E and I just don't want her do drop back down to E basically. So I want to encourage her as much as I possibly can and I think she's had a bit of a [lie bow?] moment recently. And, um, yeah. But then students like, um [Girl name], or [Boy name], or [Boy name], when they're on task, if I praise them, then I almost can see other students thinking 'Oh my goodness, like '[Girl name] is doing it right. And she's doing it. I need to do it'. Um, do you know what I mean. Because people just, I think often the kids just think that she's silly, badly behaved, can't do the maths, and actually then, when, the second they're thinking 'Oh my goodness, [Girl name] is even doing it and getting it right I need to, like, buck up'. I: OK. T: Yah, it's about, like, tapping into the perceptions of each other and stuff. T: In terms of, like, FAB 4 is when I give out the four merits at the end. So one of	--- I: OK. So, um. Oh yeah, so when the pupils are, like, working on tasks T: Mm. I: =and they have to write their answers on the white board T: Yah. I: =and when they are done, you ask them to hold them up and then you check T: Mm. I: =their work. And you kind of, like, go through all of them and say, like, 'good', 'excellent', 'OK'... like, until what point did you get, these kind of things. T: Mm. I: Why are you doing it this way? Like, why are you checking their work that way? T: Because I want to know how much they've learned. I want to know what they are up to. So some, some kids are slower at writing... so yeah getting, so I'm getting in, this is the formative feedback. --- I: Alright. And do you think, um, oh no. And, er, in generel, so this you do in front of the class. T: Mm. I: But so in general, um, when people, er, had a, when they had a test or something, would you tell them in front of the class how,	that they [T: Yeah] maybe should do a bit more. [laughs] T: Yeah.

	being the top. Um, but equally there's nothing wrong with healthy competition either and wanting to, to push on. And we do that with all the classes in all years. We also, like, publish on a board the best results from the year and everything like that. I: OK. T: So I think, we do like to publicly I: Mh-hm. T: =show them that, you know, I: Mh-hm. T: =that there is a bit of competition.	them would usually be one of those kids like [boy name] or [boy name], um, or [girl name], who just kind of sitting there and get on with it and ignore all the distractions. Um, one will usually be for a student who's made big progress with their behaviour in that lesson because they need, like, those kids usually need really fast rewards. There's, I don't, like, it doesn't usually work with those kids who are in poor habits of rewarding them five days of consistently working well. 'Cause as much as you want that to happen, you, you need to give them immediate rewards I think. I: [During the lesson you mean, yeah.] --- T: Because it basically increases the participation ratio of them, right. More of them will participate if they know that they have a chance of getting a merit at the end.	what their scores are? Like, do you speak about their performance... T: I will in general terms if, if I think 'Oh my word, you havn't prepared for this. Why on earth am I marking it?' I: Yah. T: Mm mm mm mm... But no, [absolutely] I: [Not for like specific] T: No, absolutely not. [How] I: [Why not?] T: = how cruel. Because once I got the lowest mark and the teacher read back the marks, get the, it was comprehension back, and she didn't go low enough so I didn't have my paper in. So she handed them all out again and she gave me such a bollocking in front of everyone. I would never do that.	

Table N.2

Example of Student Matrix Linking Fixed-Ability Practices and Their Motivational Beliefs

Fixed-ability practice	Student perception/motivational beliefs	Extracts
Unsophisticated task differentiation (Different tasks)	Self-concept	I: So why do you think, like, what would Madam think about you if she would always give you like easier tasks than others B: That we are not, like, as good mathematalics, mathematations... B: Stupid. B: Yeah. --- G: I don't know 'cause it might make them feel like they are more stupid than everyone else. G: Yah. --- I: Ok. So, why do you think that is? Why wouldn't she just say 'ok, that one is not doing so well so I give the easier task'? G: 'Cause sometimes like... it can knock down your confidence. […] I: Ok. I see... So you, you are saying as well that... with the... what what do you mean exactly with confidence? That it gives you more confidence if you can choose. G: Um because like, if she says 'oh you can have the easy ones because you are not good enough you are like... don't feel like you are good enough. I: Ah ok! So it would kind of give you a feeling of yeah that that your ability is not that great, you are not doing that well in class? G: Mhm hm. G: I, I think, I think that it wouldn't be, like, it wouldn't be fair. Because if you give someone, if you keep on giving someone easy tasks, they gonna find it really really really easy and then they are not really gonna, they gonna keep, like, it's not really gonna help them improve on their self-confidence. I: Hm-hm. How? Why, why? GS: Because it's like, if you keep giving them things that they find easy then I don't think they would, um, improve. And after that, if you give them like a hard task, I don't think they would understand it 'cause they've been given easy tasks all the time. I: OK. G: [whispers something inaudible].

		G: Um, yeah she said the exact same thing. But um, they actually, they will actually, like, lower, the people's with easy tasks self-esteem. 'Cause all the other people, they are, like, doing difficult tasks while they are doing an easy task. So, yah I don't think it would be really fair. --- G: Um, I was gonna say, if, if, if the teacher gave, um, like, [Pupil name] an, like, a way easier task every day, I would think it's not necessarily unfair. It's just, like, either she needs to be moved down a set because they, they're obviously, I'm not talking about [Pupil name] I: Just in general, yah. GS: =they would need to move down a set because either they're not up to that level where they can actually do it or they just need to work more, like, if they are finding it difficult, they need to understand it better. So just, pardon me, to explain the question more and then, so to help them. 'Cause if they're always on an easier one they are not actually gonna be so, we are in set A, they are not actually gonna be, like, meant to be in set A because they are working on easier questions more than the actual difficult ones.
Unsophisticated task differentiation (Different tasks)	Low teacher expectations	I: So why do you think, like, what would Madam think about you if she would always give you like easier tasks than others B: That we are not, like, as good mathematalics, mathematations... B: Stupid. B: Yeah. --- BS: Yeah, she wouldn't know how good we are if we get one an easier test and one a harder test. 'Cause if you are not as good to do the harder test then she might not think you are capable of doing that sort of thing. --- B: Yeah like... I'd be a bit annoyed because I actually know Madam knows um my ability 'cause she's um...told er my parents about my ability. So she actually... I: Oh she did? That's interesting. BS: So she actually knows my ability, so like if it's too easy I'd be a little bit annoyed if she knows how good I am and she's given me a bit easier work.
Unsophisticated task differentiation (Different tasks)	Fairness	B: [I think,] I think it would be fair because, like, I think especially, like, with someones like I mean, I'm not the best in maths, I would, I admit that, but I wouldn't be that disappointed if I got an easier task but I would be disappointed if, like, she gave me a very easy task and everybody else// --- G: Well, it would be like the same. Because say if they are, like, really struggling and then you're, like, 'cause people in my class say 'Oh that's easy' and then, like, other people might think 'Oh, shall I just say that's easy because that's what everybody else is saying?' But then if I say 'It's hard' then everyone would laugh at me. So I: OK. Would that happen you think if you would say it? GS: Maybe

Appendix O
Themes in Teacher and Student Interviews

O.1 Fixed-ability practices

O.1.1 Tasks

O.1.1.1 Unsophisticated task differentiation

According to teachers and students, none of the four teachers used unsophisticated task differentiation, such as assigning easier or fewer tasks to struggling students. Ms Evans, one of the more-fixed mindset teachers, however, provided extra tasks when students really struggled with the tasks they were given, and teachers provided extension tasks to students who finished their work early. Moreover, at Meadows School, students did “mini-tests” every day, and students reported that based on the test results, *“people get tasks based on what their weaknesses is* [sic]*.”* This was a more sophisticated kind of task differentiation than merely getting fewer or easier tasks. Other than this, students of Ms Evans and Ms Walker described that all got the same tasks but that they provided more help to those who needed it.

Ms Barns and Ms Green also did not assign struggling students with easier or less work, but they often offered tasks that were graded in different difficulty levels. Ms Barns reported that she usually wanted that all students began with the easier tasks but explained that *“if someone's struggling, they, they will stay in kind of the shallow end for the rest of the game.”* Ms Barns’ students echoed what she described, explaining, *“you start on the easy one that everyone starts on. And then if you finished that, you go on to like a harder one.”* However, other students said they could start with more difficult tasks if they felt confident. One girl even described, *“she puts on the board, if you are struggling then do this, these questions. If you, if you're finding it OK then do these questions.”* Ms Green tended to give her students complete choice over what difficulty level to work on. Yet, she remarked that students were not used to making such decisions and sometimes made the wrong one. Students confirmed this and added that she would sometimes tell them to choose more difficult tasks.

O.1.1.2 Unidimensional tasks

All four teachers taught and allowed students to use different methods to solve the same tasks although Ms Barns, one of the less-fixed mindset teachers, did not seem to actively teach different methods. Ms Evans described she would only teach methods that worked in all contexts. Her students said little about using different methods. Ms Walker reported that she would often let students choose which methods to use. However, while her students stated that they were allowed to use their own methods if they were confident enough, they said they mostly had to follow Ms Walker's way. One girl even asserted, *"everything's basically Miss' way. But then she teaches us her way but then I still do it my way."* Similarly, the boys claimed that *"Miss, um, has a way and ... if you say a different way, then she would listen to it but then she says it will confuse other people."* Ms Green described, *"I want to share um ... different methods if time allows me, then whichever method they choose, but I would tell them my preference and the reason."* Compared to the other teachers, Ms Barns was less keen about teaching different methods but students could always use a different method as long as it was mathematically correct.

All teachers seemed to make sure that no student was without work by providing extra tasks to students who finished their work early. At Meadows School, Ms Evans described that they offered a *"star question"* for students who finished the lesson starter early. She described these tasks as very rich tasks that could be solved in different ways.

O.1.2 Grouping

O.1.2.1 Group and pair work

All teachers seemed to engage their students in some pair work, but the extent to which students were allowed to discuss work with each other differed markedly. The more-fixed mindset teachers allowed much less pair work. Ms Evans and Ms Walker stated they did some pair work or tasks that partners could discuss with each other, which their students confirmed. However, especially Ms Walker's students indicated that these opportunities were rare. In contrast, both Ms Barns and Ms Green allowed their students to discuss their work with their seating partner if they wished during most of the lessons except for lesson starters and some individual work. Ms Barns was also the only teacher who offered group work, which she always did towards the end of the half term.

O.1.2.2 Seating and ability grouping

All four teachers decided the seating plan themselves rather than giving students choice. This was mostly done randomly as teachers did not know students at the beginning of the year, such as using the alphabetic order and mixing boys and girls. Over time, all but Ms Barns made some changes to the seating plan. Reasons for changes included seating stronger next to weaker students, moving weaker students towards the front of the class, and behaviour management.

All teachers were supportive of heterogeneous ability pairs. For instance, Ms Evans described that she sat some weaker students next to stronger ones. Similarly, Ms Walker described that she tended to seat students of different abilities next to each other, which her students confirmed, and that she also moved them around if they did not perform well on tests. Ms Green also reported that she preferred mixed-ability pairs, but she would not separate homogenous pairs if they worked well together and were *"challenging each other."* In contrast to other teachers, Ms Barns did not show a clear preference for either heterogeneous or homogenous pairs describing:

> *I'm wanting to put two pupils together who are a similar level so that they can pull each other along. And in some instances, I will have a stronger and weaker one because I know that that combination will work, that there is no reason why the weak one's not actually getting moving.*

Her students also noticed that both such combinations existed. Ms Barns vehemently asserted that she would *"absolutely not"* look at students SAT scores to decide the seating plan.

All but Ms Barns tended to seat students who needed more attention or distracted other students at the front of the class. Ms Evans tended to seat the students *"who struggle the most"* at the front of the class. Ms Walker also put weaker students or those with disabilities towards the front as she described *"the front row is students that I need to immediately check are doing the right thing."* Her students picked this up noticing, *"she puts you at the front of the class for the people that need, for the people that are distracted easily,"* whereas *"people who are sensible [are] somewhere at the back. So they can focus on their work."* Ms Green seated the students that distracted others towards the front rather than those that needed more help. In contrast to the other teachers, Ms Barns had no designated seats for weaker students, and students also noticed that some weaker students were actually at the back of the class.

Lastly, some of the seating arrangements were designed for behaviour management. Ms Evans seated a boy *"between two very quiet girls because he is someone who's got a bit of um more tendence* [sic] *to naughty behaviour."* Her students picked some of these aspects up as one student described, *"she puts as with people that she knows we're not gonna talk to."* One student mentioned that if they were sitting next to someone who was *"not really nice to [them],"* they could tell the teacher and she would maybe move them. Ms Evans confirmed this but also said she would usually not allow such changes. Seating for behaviour management was also a very common strategy for Ms Walker. She tried to seat students who were not well behaved next to students who were *"consistently well behaved and also not easily distracted at all."* Her students also realised that she paired students who worked well together. Ms Green also described that she used the good students to separate the chatty students. In fact, she moved the students around rather frequently to manage their behaviour. Ms Barns, again in contrast to all other teachers, thought it was unfair to make seating decisions based on behaviour management. When asked how she seated students, she responded, *"I can tell you what the incorrect way is (. ...) The one whereby the nice cooperative kids get lumbered with the poorly behaved kids."*

Students at both Meadows School and Woodpark School were set by ability. At Meadows School, the sets for maths were not only based on their maths performance but their performance across different subjects. It was only at Meadows School that the ability sets seemed to be constantly on students' minds. They very often referred to ability groups and the fact that they might move down a set even when we discussed aspects that were seemingly unrelated to ability grouping. Students at Woodpark School rarely referred to ability groups.

O.1.2.3 Stability

At Meadows School, ability sets, and therefore also the seating plans, were relatively flexible as Ms Evans explained that students *"are re-set every half term based on how they do."* At Woodpark School, the school of the less-fixed mindset teachers, students could change their ability group only after the end of the year based on their performance ranking. Ms Green was not satisfied with this procedure, saying, *"I wish they could consult us more. Yeah. Make it part of the formal, formal decision rather than purely on the score."*

O.1.3 Evaluation

O.1.3.1 Public evaluation

All teachers engaged in some public evaluation, but the extent differed greatly between classes. Ms Evans and Ms Walker, the more-fixed mindset teachers, engaged in public evaluation much more frequently and more explicitly. For instance, both asked students to share what scores they got in their weekly knowledge test in front of the whole class. Ms Evans also asked her students multiple times during a lesson to raise their hands if they got a certain number of questions correct. One of her students reported, *"she also gives out merits in front of the whole class,"* but she also gave some merits privately, such as when she checked students' books. For demerits, students described, *"she comes to you personally* *instead of saying it in front of the whole class."* One boy also described that *"she sometimes, like, points out that people have confused faces."* Ms Walker also engaged in such practices. For instance, she nominated the *"fab 4"* students at the end of every class, which was a school policy although Ms Evans did not follow this practice or at least not publicly. Students reported that Ms Walker usually nominated many different people but *"sometimes she chooses some people, like over and over again."* While she praised students very openly and very frequently, she came to students privately for criticism. When marking their books, she also handed out merits or demerits in private. During observations, her almost constant comments about students' performance or behaviour stood out markedly. Another aspect relating to social comparison was that Ms Walker frequently favourably compared the performance of this class with that of a higher set she taught, saying, *"every time we do something that 7B has done, then I tell them."* One boy actually said, *"she keeps going on about it."*

Public evaluation generally seemed very prevalent at Meadows School as Ms Evans described, *"we also, like, publish on a board the best results from the year and everything like that."* Ms Evans' classroom also had a scoreboard for aspects unrelated to achievement, such as *"participation, politeness, behaviour, equipment."* In addition, teachers could hand out marks for effort every week but these were contingent on the rest of the class because teachers *"[could] give one 10 per class, one 9 per class, three 8s, and then as many 7, 6, 5s."* The school also had regular celebration ceremonies during assemblies where teachers nominated students for *"good work and um, it's like trying the hardest and that stuff."*

The less-fixed mindset teachers seemed to engage in much less public evaluation and social comparison compared to the more-fixed mindset teachers. Ms Barns stated, *"I think I'm a lot more open about the demeriting than the meriting,"* and usually just entered merits into the system at the end of the lesson rather than doing it in front of the class. However, some of her students said she gave merits in front of the class, like most teachers at their school, which I also observed. Ms Green also awarded merits and demerits in front of the class and wrote students that were not behaving well on the board. In strong contrast to the other two teachers, Ms Barns and Ms Green mostly spoke about students' performance in general terms rather than pointing out specific students. For instance, when students worked on tasks, Ms Green commented on how students were doing, so that other students could hear as well, but she rarely referred to individual students. One boy also explained, *"she like point[s] out one person is doing good, but she wouldn't point out one person is doing bad."* Ms Barns claimed she would never read out the marks of students. She was explicitly against performance comparisons and did not want students talking about how well they were doing in front of others. The only situations I observed where students could see how the others were doing during Ms Barns' lessons, was when students solved tasks on mini-whiteboards and then had to hold them up. She then commented on how they did with this particular task. One boy also described a situation in which all students had to stand up and they could only sit down when they solved a task, so that the struggling students were the last ones standing. However, these instances seemed very rare.

O.1.3.2 Performance over process emphasis

All four teachers and their students indicated that the teachers emphasised the process and deep understanding over just being able to follow rules. They also tended to focus on effort and behaviour rather than performance in their praise and they challenged students. However, the more-fixed mindset teachers placed more emphasis on performance. This section begins with the results regarding process emphasis and then turns to performance emphasis.

The school of the more-fixed mindset teachers followed a mathematics mastery programme, a school improvement programme which aimed to develop *"mathematical reasoning"* and *"mastery, not fluency."* Ms Evans explained:

> *So fluency, if I want them to be able to multiply fractions, all I need to say to them is multiply the numerator, multiply the denominator. I could give them 15 questions on that and that would develop fluency (. ...) but mastery would be that they would understand it. They could apply it in a different context. That they could spot it in the structure of a question.*

While all teachers emphasised the importance of deep understanding, Ms Evans and Ms Green, in particular, emphasised that they wanted their students to explore the subject and not just tell them what to do and then let them practice. One of Ms Green's students described:

> *(...) when we make new topics* [sic] *she like puts (. ...) questions on the board so we just do it and like and then understand and then like the next lesson she'll go through the process with how to do it.*

Ms Barns also emphasised that she thought it was important that students *"got to try and figure out 'why?'"* In fact, she often told her students *"[d]on't follow a recipe!"* Nevertheless, she and Ms Walker always gave students worked examples and showed them exactly what they were supposed to do before setting them off to work. Teachers reported that they challenged their students and their students confirmed that they mostly felt challenged, which was particularly true for Ms Barns.

All teachers asked students to explain how they got their answers, both verbally and in writing. In particular, Ms Barns and Ms Green emphasised that it was very important how students wrote down their maths work. Ms Barns would tell her students, *"you've got to show me what you are thinking. I cannot interpret. I cannot imagine. You've got to show me,'"* and, similarly, Ms Green told her students from the start of the year, *"I'm not interested in the answer… I'm more interested, I'm interested in the working out."*

All teachers reported that they recognised effort and improvements more than performance, mostly through praise or merits. All but Ms Walker, one of the more-fixed mindset teachers, reported that they did not use much praise in general, which their students confirmed. Ms Evans said she would probably mostly praise students *"when you are surprised that they got the answer when you weren't expecting them to."* Her students recognised that Ms Evans would mostly praise improvements. One boy from her class explained, *"even if you get, like, almost everything wrong, as long as you give your best and she can see you've tried hard and maybe improve from the last time. She'll like, kind of congratulate you for trying your hardest."* Besides praising progress, Ms Walker revealed that she wanted to praise students that the class regarded *"to be like dumb almost."* Her students also mentioned that she would often praise them for progress, effort, resilience, and good behaviour. At Meadows School, students were given marks for effort every week, which were sent to students' parents. Ms Evans affirmed that she would also talk to parents about students' effort even if not required by the school.

Ms Barns also focused on effort praise, and her students confirmed that who got praised was *"not like who's the smartest, it's who has improved more."* However, Ms Barns also said that it was important that students knew how they were performing. I observed multiple situations in which Ms Green praised the class for trying hard and not giving up. She also explained that she wanted to emphasise students' resilience and that she usually awarded merits based on students' behaviour and how hard they worked. Similar to process-oriented praise, all four teachers gave process-oriented criticism such as for not completing homework, being inattentive, not working on tasks, chatting when they were not supposed to, and not bringing their equipment. An exception was Ms Barns, who did not give demerits for forgotten homework if students brought it in the next day.

I am now turning to teachers' performance and speed emphasis as well as formal and informal assessment. At Meadows School, students' performance was a great concern for both students and teachers as the school used a lot of formal assessment. In fact, there was a formal test, called the knowledge test, at the end of every week and students did daily mini-tests. Students also had to do a pre-assessment at the beginning of every half term and then a post-assessment at the end. How students performed in the latter test determined in what ability set they were placed for the next half term, which created a competitive environment. Students in that school also received specific targets they should reach. Ms Walker's students reported that she frequently referred to the post-assessment and reminded them they had to perform well on the test. Her students also often referred to these tests and even explained her behaviour with the importance of the tests. Students also described that *"they are quite pressuring around grades."* In addition, students described that *"everything is timed"* in lessons, so that they had only a set amount of time for their tasks. One girl from Ms Evans' class described, *"she's got alarms and then we can't carry that on. We have to go to the next bit so that we are on time."*

Ms Evans tended to award merits or give praise for good work, getting correct answers, and solving very difficult tasks. However, students reported this did not happen often. Ms Walker gave praise *"if someone gave a really good answer or um someone did their work exceptionally well."* Her students also reported that they sometimes received merits for getting a lot of work done. Notably, it was a school policy at Meadows School that students got a demerit if their test performance declined. Most teachers and students interpreted this as being punished for not revising or working hard enough rather than performance-oriented criticism per se. However, Ms Evans tried to give students a chance to improve by going to the school's maths clinic rather than immediately giving them a demerit. Ms Walker did seem to follow this policy more closely as a boy described, *"if they get low … it will be a demerit."* Both Ms Evans and Ms Walker used formative assessment during the lessons, but this often involved public evaluation. Ms Evans described she asked students to raise their hands to indicate how many scores they got for a given task as a type of formative assessment. In Ms Walker's class, students often worked on mini-whiteboards and had to hold up their answer for her to check their performance.

At Woodpark School, although students took maths tests, these tests did not seem to be much on teachers' or students' minds. Occasionally, they did refer to exams. For instance, Ms Barns remarked that students could not simply follow a recipe in maths because then *"when they hit exams, there is all this paper blowing around and they won't know what to do."* Similarly, both Ms Barns and Ms Green mentioned that it was important that students wrote down the working outs properly because they otherwise lost points in their exams. However, Ms Green emphasised that she wanted tests to be a learning opportunity rather than that the students *"just look at the score only."* Overall though, relatively little mention of assessments was made.

There was relatively little performance emphasis in the classrooms of the less fixed-mindset teachers although Ms Barns tended to emphasise speed. She would often ask students to work quicker or have small competitions between her and her class to see who could work faster. One boy mentioned a situation in which she pressed students to work faster, describing, *"she was like 'come one, come on' (. ...) and then she was like walking up, up to them and I don't think they knew. I think they needed a bit more time."* She also sometimes rewarded students with merits who finished tasks the fastest. Students reported that Ms Green would also sometimes ask them to work faster. Both Ms Barns and Ms Green praised students for good work or good answers such as telling them *"well done"* but neither placed a lot of emphasis on this and Ms Green even said, *"I don't like to be overpraising them."* At the end of every lesson, when students left the classroom, Ms Barns tended to let students know how they did. Her students reported she would not be disappointed if they did not do well although she would tell them *"no that's not good enough, you need to this."* Similar to Ms Walker, Ms Barns also let students do tasks on mini-whiteboards, which they then had to hold up for her to check. One of the girls in her class explained, *"I don't think it's to get it right. It's to see what we can do and what we remember."*

Most students believed that their teachers found it more important that they gave their best rather than that they performed well, although they all recognised that both aspects were important. However, students of Ms Evans' class had more mixed opinions, with some thinking she mostly cared about correct scores while others were convinced that she cared more about effort. Ms Barns' students thought she considered giving their best more important than performance although *"she wants [them] to try hard to do well."*

O.1.4 Little autonomy

O.1.4.1 Little choice

The lessons at Meadows School were highly structured and all teachers had to follow the same structure, which teachers thought allowed relatively little opportunities for students' decision making. Their students reported, *"we always do the same ... We don't know the question they're gonna ask but we know what kind of structure it's gonna be"* and *"everything is timed."* Students of both teachers reported that they had very little choice over how to work. When I asked Ms Evans' students if they sometimes had choice over what they were doing, a girl suggested maybe once since they joined the school, which was about 6 months earlier, and one girl said they were once asked what they wanted to do in the next lesson. Nevertheless, both Ms Evans and Ms Walker gave students some choice over what methods they wanted to use when solving a task as described earlier.

Ms Barns and especially Ms Green gave more choice to their students. As already described, students could often choose the difficulty level they wanted to work at. Ms Barns also let students choose with whom to work during group activities and both teachers allowed their students to work with their partners for most of the lesson if they wished to. The boy-group interviewed from Ms Barns' class did not think they got a lot of choices though. Ms Barns herself also said that while students *"were getting a lot of choice in figuring out a strategy for doing it,"* she expected them to write down their work in a certain way. As described above, Ms Green stressed that she found it important that students could choose what methods they wanted to use and one of her male students described he *"kind of"* feels as though he can make his own decisions during lessons.

O.1.4.2 Little promotion of self-regulation

Neither Ms Evans nor Ms Walker seemed to ask their students to engage in self-evaluation such as asking them to reflect on how they were doing or what they thought they should work on. One of Ms Evans' students mentioned though that she would ask them what was going on if their performance dropped significantly and another described she would sometimes ask how they were getting on *"just in case, um, she would like to give me any help."* At Meadows School, students were also given targets for their scores in tests and some online exercises they had to do. Nevertheless, students in both these classes engaged in self-assessment in their books as well as in tests. Ms Walker also reported that she frequently asked her students to assess each other's work and that they then gave feedback to each other.

Students in Ms Barns' and Ms Green's classes engaged in self-evaluation at the end of every lesson, which was school policy although the students reported that not all teachers did it as regularly. Students had to indicate how they thought they did during the lesson and what they needed or wanted to work on more. One boy explained, *"you wouldn't put like, 'I'm happy with this' because she doesn't want that. She wants us to put a reason why we're happy with it."* Ms Barns' students also reported that she asked them at the beginning of the year *"where do you think you are with maths?"* I cannot tell whether students got targets for their test performance as neither teachers nor students mentioned this. This indicates that even if there would be targets, these did not seem to have great importance. Similar to the other two teachers, both Ms Barns and Ms Green asked students to engage in self- and peer assessment.

O.1.5 Differential value

All four teachers and their students reported that teachers tried to ensure that all students participated in the classes, not just the high performers. There were thus no marked differences between teachers. All teachers seemed to especially often call those students who were weaker in maths, paid less attention or were shy and did not volunteer answers. For instance, students of Ms Evans' class reported that she picked those that usually did not talk much. Ms Evans also picked students who did not raise their hands and overall her students thought that *"normally she chooses, she chooses lots of different people,"* and showed an equal interest in all students. She might only be a bit *"biased"* against students who did not behave well in a given lesson but then *"she kind of refreshes it, like, in the next lesson."* Ms Walker's students also thought that they were being picked equally but maybe she called more often on the weaker students or those that were not concentrating. Similarly, Ms Barns tried to *"make a point of picking the kids who don't put up their hands"* or those who were not participating as much. When students worked in pairs, she also checked that all students were engaged rather than that it was just one student doing all the work. Ms Green also tried to encourage all students to participate, especially those that were not paying attention, and she also called on students who did not raise their hands, which her students confirmed. Her students also reported that she picked students quite equally although she usually went for *"the people that aren't as, like, good at the subject."* Overall, they thought that Ms Green gave equal attention to all students except to some that were always interrupting the lessons.

Nevertheless, teachers were also aware that they did not always manage to call on all students equally. Ms Walker said, *"there are definitely times where I think 'oh my goodness, I haven't asked that kid for ages,'"* while she also did not think *"there's anyone who I completely ignore or leave out."* Likewise, Ms Evans described, *"sometimes I get to the end of a lesson and think 'who hasn't answered a question yet,'"* but she believed that she probably did not call on all students equally. At the same time, Ms Evans admitted:

> *(...) there are some kids who, if I do, if I want them to get the an-, if I want to be quick about it then I might ask a question saying, like, oh I know you are going to get it right.*

Hence, she picked the high performers to move on the lesson more quickly, which her students noted. This was not unusual for the other teachers either. Ms Barns also described that *"as the discussion gets harder, fewer pupils are kind of in the ring to be able to answer* *the questions."* Her students reported that she usually picked students who did not know the answers but if many tried and failed, she had specific students she called on who usually got it right. One boy reported that he experienced a lesson where most students struggled; he described that Ms Barns *"just goes straight away to the person who knows it."* Similarly, Ms Green said when I asked if she called all students equally, *"I'm trying to, but I know I didn't manage to because some questions I know it's hard. Only the good ones can answer. So, for those question I would call them more often than the other one."* She did say though if weaker students wanted to answer a difficult question, she would pick them first.

O.1.6 Little promotion of risk-taking

O.1.6.1 Little encouragement of help-seeking

All four teachers encouraged their students to seek help from them, and they regularly checked whether students understood what they explained or the tasks they worked on in case they needed help. It seemed that teachers preferred to give students individual support rather than to discuss questions on the board although they tended to repeat something on the board if multiple students had the same question. The teachers also seemed to allow that students came to them at the end of the lesson to seek help. While students were working on tasks, all teachers walked around the classroom to check if they needed help, although this was less frequently observed for Ms Barns. Ms Evans reported that she had some students she liked to check on more regularly. However, one of her students explained, *"she doesn't actually, like, pick on the people who she thinks, they put their own hands up and they also help."* Similarly, it seemed as though Ms Green offered more help to struggling students as one girl described, *"she comes [a]round tables to help people which she knows maybe like struggle a little bit."*

There were also instances in which students did not feel as though they received the support they needed. Some boys of Ms Evans' class complained that when they did not get many correct scores, she *"doesn't go to all of them independently and helps them."* Most students felt as though they could ask Ms Evans for help, although one student admitted she generally did not like to ask questions to teachers in that school. While Ms Walker mostly encouraged students to ask questions, she also said students often asked questions because they did not listen and then she would go up to them and tell them *"actually, when you asked that question, it wastes our time because you clearly weren't listening."* This was echoed by one of her students describing, *"half the time she gets, like, quite angry because, um, she thinks that we are attention seeking...but then we just don't really understand it."* Moreover, some of her girls revealed they did not always get her support. Similarly, while Ms Barns reported that she generally checked if everybody understood, she said if *"they've been sitting there and sleeping. I'm not impressed."* Students' opinions regarding how much she helped differed. For instance, one boy described, *"I was finding it a bit hard today and she came around and helped me,"* to which another replied, *"[s]he never helps you."* A girl from her class described that she even suggested them extra resources, such as websites, whereas a boy maintained that she just answered their question but did not really provide help. Most students felt comfortable to ask Ms Barns for help, but two said that they did not feel comfortable, especially if she had been explaining something for a very long time. They generally worried about asking Ms Barns for help although they never received a negative reaction from her. All of Ms Green's students felt as though they could ask her for help.

One discernible difference in the encouragement of help-seeking was that the more-fixed mindset teachers did not encourage much peer support while Ms Barns and Ms Green did. One boy from Ms Evans' class pointed out that they could only ask their peers for help *"if she tells you to ask them (. ...) if you just help them for no reason, then you get a demerit."* The girls also reported, *"they don't really allow to help us, like, the people who sit next to us. We only have to do the work by ourselves."* This was although both teachers reported they liked to seat stronger and weaker students together, so that the former could help the other, and some students reported that they can turn to their neighbours for help. These somewhat conflicting messages indicate that there were likely some designated times or activities during which students could talk to each other, while they had to work alone most of the time. For instance, Ms Walker reported that she did not allow a lot of partner discussions towards the end of the half term when students needed to revise. In contrast, both Ms Barns and Ms Green allowed students to work in pairs most of the time and also encouraged them actively to help each other. Their students reported that they could help each other as long as they worked with their partner rather than with other students around them.

O.1.6.2 Unconstructive handling of mistakes

All students thought their teachers did not mind mistakes, indicating that the teachers showed mostly constructive reactions. Handling of mistakes did not differ markedly between teachers although only students of Ms Evans and Ms Walker described that their teachers reacted happily to their mistakes, while those of the other two teachers said they did not *"make a great deal out of it."* At Meadows School, teachers were supposed to celebrate mistakes as Ms Evans explained, *"a scriptive thing that we are trying to say as much as possible is 'I'm* *so glad you've made that mistake' or 'thank you for making that mistake because it brought out a really important point.'"* Ms Walker echoed this.

All teachers usually discussed mistakes and showed students where they went wrong. Ms Evans often explained mistakes and, similarly, I observed multiple instances during which Ms Walker held up mistakes to discuss them. Ms Barns also described, *"I will u-, usually try and find something positive going on. Um, or draw as to 'now why are you thinking that' and sort of use that as a starting point to, for the discussion"* and one of her students recalled, *"she said like if we don't make mistakes, we can't learn from them."* Her students confirmed that she usually explained mistakes. Ms Green described that she learned about a research project at a school she used to teach at, where they looked at how teachers can help *"students to make lots of, not helping them make mistake but helping them to learn it through their mistakes,"* which was an approach she really liked. There were also instances for all four teachers in which they gave students more time or another chance to solve a task. For instance, one of Ms Evans' students described, *"sometimes instead of just moving on to another person, she'll say, 'no, that's wrong. How about, you're thinking about that. Try again thinking of that.'"* Similarly, a student of Ms Barns described, *"she just wait[s] a couple of minutes"* when someone made a mistake and *"if she can see something wrong she's like 'how can you improve that?'"* One of Ms Green's students explained that if someone made a mistake, *"she almost makes you go back and think about it again."*

All teachers also showed some unconstructive handling of mistakes. For instance, if many students tried but could not solve a task, the teachers would sometimes provide the correct answer themselves, although Ms Green claimed she tried to avoid this. A common teacher reaction was to call on another student after a mistake occurred. Ms Walker described that she was critical of mistakes that occurred due to a lack of effort rather than a lack of understanding. Similarly, I observed one situation in which Ms Green was rather angry with a student who made a mistake, but both she and her students explained that this was because this student never listened and thus gave absurd answers. Moreover, although Ms Barns said she *"certainly wouldn't rubbish [mistakes],"* she described that she would sometimes tease students about mistakes or react with an *"Oh really?"* if the student could take it, arguing that *"it's good-natured."* One boy even reported that she sometimes seemed disappointed when they made a mistake although the other students did not agree.

O.2 Motivational beliefs

O.2.1 Mindset

O.2.1.1 Teacher mindset

Despite not having completely rejected a fixed mindset in her questionnaire responses, Ms Evans indicated during interviews that students could improve their performance and that she wanted to give them ample opportunities to do so. This became particularly apparent in her going against the usual school policy of giving students a demerit for declining test performance. Instead, she gave them the opportunity to attend the school's maths clinic, where older students teach younger ones during school breaks. One boy confirmed this, describing:

> *(…) she's more aiming at your education than your behaviour to be honest. Because she wants you to improve on your intelligent in maths (. …) let's say for maths clinic, she doesn't give you a demerit to feel bad. She gives you opportunity to improve. So you can do it better next time.*

Discussions with Ms Walker showed that she held some fixed- and some growth-mindset beliefs. For instance, she did not believe that many of her students in the lower ability set could move upwards as she thought they were not in that set because of their behaviour. Specifically, she explained:

> *(…) there are really few students in this school actually who are in the lowest sets based on behaviour. But the ones that are always know. Like they, like [girl] for example, knows that if she consistently had worked harder and if she was always behaving and always in her lesson, she probably wouldn't be in D.*

However, she also stated that even though people had different abilities, people could learn if they *"put in the hard work."* She also mentioned that many students thought that making mistakes was a sign of being *"stupid,"* which she wanted to interfere with, indicating her disagreement with such a fixed-mindset belief.

Ms Barns, who completely rejected a fixed mindset in the questionnaire, showed both growth- and fixed-mindset beliefs in interviews. A notable instance was where she stated, *"the fact that someone is on a Level 3 and someone else is on a 5, well that's just what it is. And, and the children know. From when they start school, they know in f-, their pecking order."* Especially the fact that she said this was the case from the start of school points towards a fixed mindset. Only when I asked whether she thought students could change the order, did she say, *"Mm! Absolutely! (. ...) If they suddenly sort of start joining the dots, yes of course!"* Nonetheless, her previous statement was more in line with a fixed mindset and she referred to an almost miraculous improvement rather than improvement through effort and better learning strategies. More in line with a growth mindset, she later explained that she did not want to pre-judge what students could and could not do. Similarly, she did not believe that primary school performance was a good indicator of students' potential because *"they can be schooled hugely."* In addition, she did not only apply a growth mindset to her students but also recognised that she was a much better teacher now compared to when she started.

There were only a few instances where Ms Green displayed her mindset beliefs, but when she did, she exposed both growth- and fixed-mindset thinking contrary to her questionnaire responses. For instance, she stated, *"[s]et 4 is unlikely they're going to choose A-level maths. Um, they might use basic like um, calculate percentage of um money, and that's probably what's people experience in the daily life."* Given that these students were only in Year 7, there should be plenty of time for them to improve, which is why her thinking here points towards a fixed mindset. However, she also stressed that she aimed to *"develop their maths thinking skills."*

O.2.1.2 Student mindset

Ms Evans' students believed that they could improve by doing extra work or getting help. One girl also pointed out that making mistakes and then correcting them or asking a teacher helped her to improve. And another girl described that students could improve if they got challenged but not if they got easy tasks all the time. Two girls from Ms Walker's class actually said that they learned about a growth and fixed mindset during a school assembly. No other students participating in interviews indicated that they had been introduced to the concept. One of them explained:

> *(...) apparently everyone has to be a, er, in a growth mindset at like all times when we are learning. Because it, I would say if we, a fixed mindset is like 'oh I can't do this. I wasn't born like this.' Yah. 'Other people are better at maths than me 'cause they were born with, like that.'*

These girls, as well as the two boys I interviewed from Ms Walker's class, indicated multiple times that they believed they could get better if they revised.

Ms Barns' students also indicated that they could get better if they worked hard. One girl stated, *"as you gradually grow you can improve."* Ms Green's students also showed a growth mindset. For example, one girl argued that if *"you are not trying that hard that means you are just always going to stay at the level that you are. And but if you try the harder ones that means you gonna learn more and move upwards."* Similar to the teacher interview, there were only few such instances.

O.2.2 Achievement goal orientation

O.2.2.1 Teacher goal orientation

Only Ms Green revealed some of her teaching goal orientations, although this was extremely rare. In fact, there was only one instance in which she expressed a performance-avoidance goal. She described that some students complained about her to others and that she would wish they would go to her directly. She criticised, *"they don't come to the teacher, they tell their parents, they tell their head of year. All the way through the teacher look like idiot!"* She emphasised though that she was not the only teacher in that school feeling this way.

O.2.2.2 Student goal orientation

Students from all four teachers expressed mastery goals. They indicated that they wanted to increase their performance, push their brains, and be challenged because that made them think. For instance, a boy from Ms Walker's class announced, *"I don't like easy things (...) Because, er, because if I already know it then no point doing it."* Similarly, students in Ms Green's class said it would be unfair to get easy tasks because then they would not be learning, and they liked that Ms Green offered different difficulty levels because *"if you wanna challenge yourself go to the ... straight to the high ones."* Students in Ms Barns' class did not express their goal orientations often, but one girl stated that she welcomed performance feedback from her teacher because that helped her to know how to improve. She further explained:

> *(...) if she said I wasn't [doing well], that means I know that I need to do more stuff to make it better. If she didn't tell me, I, it would be like, how am I meant to make it better if I don't know?*

Ms Evans' students expressed much more performance goals compared to the other three classes. This was often linked to Ms Evans' praise and public evaluation. One girl described, *"you want her to praise you even more and so, like, you work even harder,"* and a boy said that he felt sad if he did not get *"gratitude"* for his good performance. When I mentioned to students that she often seemed to ask how many scores they got on a task or test, one boy exclaimed, *"Yeah! I love where Miss asks that! Like, personally I love it. Because then everyone knows I got it right!"* One boy added that it can actually be quite humiliating for those who did not perform well, and girls also revealed they did not want others to know that they struggled. Students recognised that people reacted quite differently to performance comparisons, with some working even harder to outperform others while some *"feel really like, um, bad about it and then they will like coy, like doing anything to do with maths."* One boy described, *"there is like a mixture of, some people will be, like the people with the glass half full will start kind of working towards it and then the half-empty people will probably, like, give up on it."* He further explained that the latter people were scared to give answers in front of the class. All students described how this led to cheating behaviour because people did not want others to know how poorly they did. Some students also expressed they did not want to appear stupid in front of the teacher, and thus avoided asking questions, but this was less prevalent than appearance concerns for their peers. Only one girl explicitly said that she had no performance appearance concerns.

In contrast to Ms Evans' class, there were no instances of performance-approach goals in interviews with Ms Walker's students. However, one girl revealed that there were some students with performance-avoidance goals, explaining:

> *(...) 'cause people in my class say, 'oh that's easy' and then, like, other people might think, 'oh, shall I just say that's easy because that's what everybody else is saying?' But then if I say, 'it's hard,' then everyone would laugh at me.*

Ms Walker recognised that there were some students with performance-avoidance goals in her class while other students did not mind as much what their peers thought. She explained, *"you can rely on the students who are happy to share their mistakes and ask questions to answ-, help answer questions of those who aren't gonna ask the question."* She also noted that many students were embarrassed when making mistakes. She said this was also why some students cheated at the beginning of the year.

There were very few instances in which Ms Barns' students expressed their goal orientations. One girl described that she did not like to ask questions in front of the class because she struggled a lot, indicating she had a performance-avoidance goal. So overall, in this class of this less-fixed mindset teacher, no type of goal orientation became very apparent. Similarly, there were hardly any instances of performance goals amongst Ms Green's students. Only one boy admitted he always wanted to be one step ahead of his partner, showing a performance-approach goal. There were no instances of performance-avoidance goals. However, many students in this class described cheating behaviours; Ms Green usually put the correct solutions on the back of worksheets, which some students then copied. The major difference to the cheating behaviour in Ms Evans' class, however, was that this cheating was mostly motivated by laziness rather than performance-appearance concerns. Overall then, Ms Green's students displayed a lot of mastery goals and almost no performance goals.

O.2.3 Attribution

O.2.3.1 Teacher attribution

Ms Evans offered very few explanations for students' performance but when she did, she made effort attributions. For instance, she argued that test performance could show you how much effort students invested in learning. The other three teachers made both effort and ability attributions. As already described, Ms Walker assumed that most of her students had a too low ability to be in a higher ability set, while only a few could move upwards if they worked harder. She also said that performance was often due to effort or how distracted students were in lower-ability classes, implying that effort matters less in higher-ability sets. Nevertheless, she believed that students could learn if they worked hard despite having slightly different abilities. Lastly, she made the situational attribution that sometimes, poor test performance could be explained by the students not having had the right partner during lessons. Ms Barns, a less-fixed mindset teacher, also showed a combination of effort and ability attributions. She explained that some people got the work done because they were more intuitive whilst others worked very hard. In general, she showed that she believed that effort could cause better performance. She also attributed getting work done relatively quickly to students' effort and concentration rather than their ability.

Ms Green also made ability and effort attributions, and to a smaller extent also situation attributions. For instance, she attributed one student's performance to his high capability and also mentioned that sometimes students overestimated their capability and chose tasks that were too difficult for them. She also distinguished between different ability sets by stating, *"if that's Set 1 probably most people get it very quickly. But Set 4 they ah take a long time to get this stuff."* On the other hand, she also said that some people could actually do more difficult tasks than they chose but that they were lazy. She also believed that if students really wanted to try, they could solve very difficult tasks. Ms Green sometimes explained students' mistakes by her having set too difficult tasks or not having explained something properly. And she also attributed their difficulties to the primary school curriculum.

O.2.3.2 Student attribution

Ms Evans' students clearly showed that they thought that both effort and ability explained performance. They thought that the more ability you have, the less you needed to work, pointing towards a reverse relationship between the two. One boy justified that Ms Evans tended to praise the improvers more, explaining:

> *(…) they have learned more. Because she understands when they started with and how they ended with. Maybe she knows that you are a bit more cleverer* [sic] *and that, er, she knew you would be able to push yourself to that.*

The students also acknowledged, *"everyone learns at a different pace and we learn different things in a certain amount of time."* One boy even specifically stated that some students got few correct answers because they were *"dumb."* Students also said they could exactly pick out the smartest in class because those got the answers correct. When one girl remarked, *"they always revise,"* another quickly countered, *"No, I know some that don't revise at all and they just get it right."* All students believed though that they could improve their performance through effort. Lastly, some students explained that the situation could sometimes prevent them from being able to learn everything, such as long school hours, too much homework, stress, or the teacher not explaining something properly.

Ms Walker's students' attributions often became apparent in how they interpreted her praise or criticism. For instance, students described that they got a demerit for a lower score on a test because it showed that they had not revised. And while one of her girls interpreted receiving merits for getting work done as being praised for being fast, another suggested that maybe those people concentrated more. Both girls believed though that you could get better when you put in the work and both thought that performance was not due to what you were born with but depended on what you learned, referring to what they heard about mindsets during a school assembly.

Ms Barns' students also thought that students had different abilities but that they could still improve through effort. One girl explained:

> *(…) you might be born like maybe not being as good as someone else's, wait maybe not born with it but just how you learn when you are first learning stuff and you might not get it but um as you gradually grow you can improve that. You just, it's just about working hard or not.*

Similar to Ms Barns, her students also attributed finishing a task quickly to having concentrated well and worked hard. Yet, one girl stated that their SAT results showed that the students in their class all had a similar capability, and another girl noted that some people were just faster at calculating than others, so that even if all put effort in, some will be better.

Ms Green's students echoed her belief that both ability and effort explained performance. They said, for instance, that some students were just better than others, that students had different abilities and learned at different paces. But they also acknowledged that paying attention was important and that some students just loved maths while others did not want to try. One boy pointed out, *"even if you are not that great but then you are paying attention you could be better than somebody who's better than you and that isn't paying attention."*

O.2.4 Student self-concept

Here, only the general self-concept of students, unrelated to FAPs, will be described. This section is very brief because only few instances occurred in which students commented on how confident they were with their maths work. Many girls tended to portrait themselves as struggling or they doubted themselves in maths. For instance, one girl from Ms Evans' class said, *"sometimes I feel like my questions are a bit stupid."* Another girl described, *"[Girl] is, like, quite clever, she is more clever* [sic] *than me (. ...) I only do about two and she does about, like, 12 questions. And I'm just like 'what is wrong with me?'"* She further explained that she was always astonished by how much work others did. Some other girls described themselves as being *"not that good at maths"* or stated they *"struggle a lot."* Girls from Ms Barns' and Ms Green's classes also described that there were students who did not feel confident about their maths work and that these students picked easier tasks. There were also some students, mostly boys, who demonstrated a strong self-concept. Boys from Ms Walker claimed they were *"excellent," "fantastic,"* and that things were *"easy"* for them, and one boy from Ms Green's class also said that he was very capable. However, one boy also described that a girl in his class was very sure of herself as he said she often turned around to him saying things like, *"she is so smart and, like, I'm so stupid. Like, how can stupid and smart one be together."* Moreover, one girl in Ms Green's class stated, *"me and my partner we are like about the same level with each other and we know most things."*

O.2.5 Teachers' perceived ability and expectations for students

Ms Evans emphasised that she did not want to communicate differential expectations but instead wanted to show high expectations to all students. However, she did not expect the same of all students. For instance, she explained, *"I wouldn't expect them all to finish in the time."* She also had students she held higher expectations for than others, pointing out some students she always expected to reach the extension tasks. Similarly, she said that it was sometimes difficult to challenge all students because *"there are some that are very very very capable. Like [student], is very very capable."* She also stated that she tended to praise students who solved tasks that she did not expect them to solve, which was something that her students picked up. Similarly, the students described that Ms Evans would not pick certain students because she knew that the question was too easy for them. Students thought that Ms Evans had high expectations for them in general or for them personally, which one boy explained with them being in the highest ability set.

There were only a few instances in which Ms Walker expressed perceived ability or expectations for her students. She perceived some students to have a higher ability than others describing, *"[Girl], when she first came into D, came from set E and was much lower ability than the rest. Um, whereas [Boy] was slightly higher attaining than the rest of the class."* And again, given that she thought most students were in her lower set because of ability rather than behaviour, this points to relatively low perceived ability overall. The boys in her class said she had high expectations for them and that *"she has faith in us."*

Ms Barns had very high expectations for all her students but explained that what *"doing well"* meant was different for each student. However, she expected everybody to work very hard and stated, *"we've got to challenge them."* Ms Barns' students confirmed that she expected a lot of them but also that she was optimistic. Students did not seem to perceive differential expectations.

Ms Green recognised that there were relatively big differences in how students performed although her students were in the same lower-middle ability set. Yet, surprisingly, she had relatively low expectations for this set in general. For instance, she did not expect them to do A-level maths and thought that they will only need quite basic maths skills in their lives. She also seemed to have different standards for different students as she explained that she wanted to have a high standard for a boy who she perceived as very capable. In addition, she said that *"only the good ones can answer"* very difficult questions and stated she knew *"the ones always having difficulty for everything I teach on the board."* In contrast to other teachers, Ms Green even told students explicitly that she did not expect that all students were able to use a given method. Students also said that they knew who the teacher expected to be able to use the method. Despite all of this, Ms Green herself emphasised that she had very high expectations and that she did not want to lower her level to theirs, especially because their level was changing as she kept on trying to stretch them.

O.2.6 Nature of maths

Ms Evans described maths as a complex subject in which many concepts and ideas are connected. To understand these connections and gain deep understanding, she said students could not simply be told what to do and then practice the rules. She declared she wanted to develop students' map of maths, describing:

> *(...) you know we are talking sometimes in maths about, like, the difference between knowing how to get from A to B and A to C, and A to D, and A to E, and knowing the whole map. So, like, a town that you, a town that you visit on holiday, you might just know how to get from the hotel to the bank and from the hotel to the restaurant. A town that you know really well, you could get from anywhere to anywhere.*

The two less-fixed mindset teachers shared this view as they both emphasised that maths was not about memorising. Ms Barns emphasised that *"maths isn't a recipe"* and that students needed to understand why they were doing what they were doing, explaining *"if they just learn how to do things, it's like, their maths is like a paper trail."* Similarly, Ms Green stressed that maths was not about memorising and that she wanted students to figure things out by themselves. Although Ms Walker also emphasised the importance of conceptual understanding, she did not think that exploration was helpful in maths but that students should be provided with all the knowledge they needed before setting them off to work.

Ms Walker and Ms Barns both addressed the enjoyment of maths, but they had different ideas about what students derived enjoyment from. Ms Walker described that students enjoyed maths when they were able to do it, especially when they were quick. In contrast, Ms Barns described repeatedly that maths was a difficult subject because *"we are teaching them to think"* and that she also let students know that maths was supposed to be hard. She frequently mentioned that engagement and enjoyment in maths were important while she clarified:

> *I'm not trying to entertain them. I do want my lessons to be interesting that, that I think is my responsibility, that I have to teach them good maths and it should be interesting. If it's mainly c-, cerebral, well, that's the nature of the beast.*

Ms Barns viewed maths as linear, something that builds on something that came before and said, *"there is nothing more boring and demotivating that going around the same ellipse year after year."*

O.2.7 Student emotion

When students were asked what they thought about certain teaching practices, they often responded by describing their emotional responses. Here, I merely provide an overview of the kind of emotions that students expressed. Positive emotions included feeling good or considering something to be good (*"you feel really good about it"*), being happy (*"they'd be happy"*), and feeling proud (*"they are gonna feel happy and proud of themselves"*). Negative emotions were more varied than positive ones and included being upset (*"making other peo-, children feel upset"*), feeling bad (*"I feel horrible"*), not liking something *("It puts you off"*), as well as feeling disappointed (*"it's just really disappointing"*), humiliated (*"they feel like really humiliated"*), and embarrassed (*"It's a bit embarrassing"*). It also included being scared or worried *("I was scared to ask my teacher"*), angry (*"I don't know why but I have this feeling as angry"*), stressed or pressured (*"it's really stressful"*), bored (*"I kind of find it boring"*), annoyed (*"you would get really annoyed at the teacher"*), and even feeling unworthy (*"I wouldn't really feel like I'm, like, worth her time"*). Another aspect of this theme are judgements about the fairness of different teaching practices. Students frequently commented on whether they regarded something their teacher did as fair (*"I think that is actually fair"*) or unfair (*"I kind of found it kind of unfair"*). Lastly, there were also many instances in which students felt indifferent towards a practice or did not mind or care about it. Such comments included *"I don't mind"* and *"It's not a big deal."* Such indifference was particularly common for the two boys interviewed from Ms Walker's class.