WILEY

Measuring morality in infancy: A scoping methodological review

Jennifer Lavoie[1] [iD] | Aja L. Murray[2] | Guy Skinner[3] |
Emilia Janiczek[2]

[1]Moray House School of Education & Sport, University of Edinburgh, Edinburgh, UK

[2]Department of Psychology, University of Edinburgh, Edinburgh, UK

[3]Institute of Criminology, University of Cambridge, Cambridge, UK

Correspondence
Jennifer Lavoie, Moray House School of Education & Sport, University of Edinburgh, Holyrod Rd, Edinburgh, EH8 8AQ, UK.
Email: jennifer.lavoie@ed.ac.uk

Abstract

We conducted a scoping review of methods that have been used to measure infants' (under age 2) moral development. We aimed to assess the state of knowledge through a thematic overview of the methodologies that have been used and the specific constructs studied. We found that the majority of studies used an experimental methodology, and within this, infants' actual behaviour and their evaluations were the most common sources of information. An evidence map depicting concept delineation between studies and presenting concepts as related to an underlying moral sense, as prosocial (emotion and behaviour), and as antisocial components (emotion and behaviour) is provided. Just under one-third of studies were longitudinal, and a high percentage reported a statistically significant longitudinal relation for moral development. Results highlight a need for measures that can be used longitudinally at different stages of development so that trajectories can be observed and mapped to behavioural outcomes, such as conduct problems.

KEYWORDS
infancy, moral development, moral reasoning

1 | INTRODUCTION

One of the hallmarks of early social development is the emergence of the ability to reason about choices that have a moral component, for example, obeying the rules. Infants demonstrate the ability to begin reasoning about actions that have moral underpinnings at a very early age, within the first year of life (e.g., Hamlin, Wynn, & Bloom, 2007; Jin, Houston, Baillargeon, Groh, & Roisman, 2018; Meristo & Surian, 2014), and they increasingly continue to demonstrate the ability to reason about complex situations in which they must juggle multiple social and moral rules as they mature (Smetana, Metzger, Gettman, & Campione-Barr, 2006). With the increase of the study of morality (e.g., see reviews Dahl & Killen, 2018; Hamlin, 2013; Lapsley & Carlo, 2014) also comes a myriad of methods for measuring morality in infancy, as well as challenges for aligning definitions and operationalizations of morality in such a way as to optimally advance the field with future studies.

Illuminating infants' emerging sense of morality is of considerable value for understanding the early foundation of moral behaviour and how this develops over time. For example, the foundation of early moral choices and behaviour may set the stage for developmental patterns through childhood and adolescence, and may explain how some children tend towards a trajectory of morally and socially accepted behaviour from an early age, while others tend towards trajectories of unaccepted behaviour, such as delinquency and violence. As of yet, this has been neglected as a focus of empirical research, despite the considerable value and ability of early social behaviour (e.g., aggression) to predict later social behaviour (Campbell, 1991; Loeber, 1991; Moffitt & Caspi, 2001; Patterson, DeGarmo, & Knutson, 2000; Shaw, Bell, & Gilliom, 2000; White, Moffit, Earls, & Robins, 1990).

To support such research, there is value in first conducting an overview of existing studies that have focused on the early moral development of infants (2 years of age and under). We used a scoping review technique (Arksey & O'Malley, 2005; Colquhoun et al., 2014), following the guidelines of the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR). Scoping reviews are used to identify and synthesize a broad body of literature that is not focused on a specific set of questions. For example, a scoping review tends to be useful for 'answering broad questions, such as "What information has been presented on this topic in the literature"' (Sucharew & Macaluso, 2019, p. 416). Scoping reviews are well-suited for large bodies of literature that are heterogeneous (Peters et al., 2015), and can be useful for assessing the scope of research literature (Grant & Booth, 2009). This approach can also be useful for revealing themes and identifying gaps in the literature (Peters et al., 2015).

Our scoping review surveyed studies of moral development in infancy to assess (a) the methods that have been used to measure early moral development, (b) to characterize the methods that have been used and how these map onto domains of moral development, and (c) to potentially identify early flags for conduct problems within the moral domain of infants' behaviour. The findings of our scoping review can be used to assess the current state of the evidence, and to inform future study designs in the field.

1.1 | Defining morality

Morality is a broad term and it is thus important to consider how it has been and can be defined. Infants' early moral development provides a foundation for how they engage in society in terms of right (or accepted) versus wrong (unaccepted, against the rules or laws) behaviour. Morality is not defined in a consistent manner in the field; a situation complicated by arguments that what is moral for one may not be moral for another (Greene, 2007; Turiel, 2015; Wynn & Bloom, 2014). Multiple perspectives overlap, with perhaps key divergences on where a sense of morality originates and whether it has a stable or dynamic developmental process. For example, some researchers argue that morality is intuitive and that infants have an innate social core (e.g., Hamlin, 2013; Van de Vondervoort & Hamlin, 2016). Some researchers consider morality to be more deriving from a developmental process that occurs from exposure to culture and rational thinking (e.g., Thompson, 2012; Wynn et al., 2018). Others view morality from a constructivist approach that considers morality is constructed as part of a developmental process (Dahl & Killen, 2018). This latter perspective views morality from a focus on how one's actions affect third parties, or, in other terms,

'prescriptive norms concerning others' welfare, rights, fairness, and justice' (Dahl & Killen, 2018, p. 2; as used by Killen & Rutland, 2011; Turiel, 2015). We adopt this latter definition as the framework for our review and expand on this perspective to include the element that morality stems from an internal standard, which becomes the basis that determines right actions from wrong actions, with such a standard being reflected in part through external rules that act as social regulators for the benefit and protection of society. We highlight that this definition is not specific or dependent on age or developmental abilities, but rather it represents a conceptual umbrella definition for morality as children's moral development is a gradual progress from early in infancy onwards through adulthood.

Operationalizing morality. Morality as a construct can be divided into three subconstructs that are more easily operationalized in child and infant-based research. The first is moral cognition or moral judgement, which is the ability to reason or think about value-based decisions (Buon, Seara-Cardoso, & Viding, 2016). The ability to measure moral cognition is increasing due to advances in the field of neuroscience (Narvaez & Vaydich, 2008), and one example of how this has affected the ability to study moral cognition is through the use of eye-tracking in infancy studies (e.g., Cowell & Decety, 2015). In young preverbal infants, looking time, as a proxy for distinguishing between expected and unexpected events following the violation of expectation (VoE) approach (e.g., Hamlin et al., 2007; He, Bolz, & Baillargeon, 2011; Hespos & Baillargeon, 2008; Jin et al., 2018) has been used as a proxy for moral cognition (e.g., Meristo, Strid, & Surian, 2016). However, the use of this approach has also been criticized as not necessarily representing true infant preferences as a proxy of their moral thinking on a subject (Tafreshi, Thompson, & Racine, 2014).

The second subconstruct within morality is moral emotions. Moral emotions are emotions that arise based on evaluations of how one's actions may affect another, as well as those that arise based on evaluations of another's emotional state (Malti & Dys, 2015). Following from our above-expanded definition of morality, moral emotions may also arise based on violations of one's own internal standard. Moral emotions include those such as guilt, shame, remorse, compassion, gratitude, and elevation (Malti & Dys, 2015; Tangney et al., 2007), and can be connected to one's own feelings or feelings for another.

The third subconstruct within morality is moral behaviour or action, which is manifest behaviour that has a moral foundation or component to it, for example, honesty, obeying the rules, and treating others in socially accepted and ethically correct ways (Krettenauer, 2012). Many studies that have focused on the early emergence of morality, broadly speaking have studied early social behaviour that is thought to be a precursor to early moral judgement abilities (Dahl & Killen, 2018). This can be seen in, for example, seminal works by Warneken and Tomasello (2006) who studied spontaneous helping behaviour in toddlers. Behaviour in the moral domain has a value-based component, and often impacts the self or another person (e.g., behaviour i.e., considered right or wrong, and behaviour that can harm the self or another person; Smetana, 2013), whereas social behaviour is largely based on social conventions or personal choice (Killen & Smetana, 2015). However, given that many studies have relied on measures of social behaviour as a proxy or precursor to moral judgement, this systematic and scoping review included both measures of moral or social behaviour with a moral component (e.g., sharing and comforting) to ensure that overlap was captured.

Although these constructs (moral cognition, emotions, and behaviour) are correlated, the correlations are far from unity, suggesting that they are not only conceptually but also empirically distinct (e.g., Krettenauer, 2012; Lakoff & Johnson, 1999; Lavoie, Nagar, & Talwar, 2017; Lavoie, Leduc, Crossman, & Talwar, 2016). This distinctness can also be observed at a developmental level where many of the contributing developmental processes are maturing at different rates (Krettenauer, 2012).

1.2 | Development of early morality

Evidence of infants' morality can be detected early. It can be seen, for example, in their general preferences for prosocial characters who help others even as young infants, rather than antisocial characters who work against others' efforts (Dunfield & Kuhlmeier, 2010; Hamlin et al., 2007; Jin & Baillargeon, 2017; Meristo & Surian, 2014; Olson & Spelke, 2008; Ziv & Sommerville, 2017). It can also be seen in their tendency to spontaneously help adults around them early in their second year of life (Warneken & Tomasello, 2006), and by their desire to repair the harm

that they have accidentally caused to another person (Hepach, Vaish, & Tomasello, 2017). Young children's sense of morality can also manifest in ways such as obeying the rules placed on them by external control agents (e.g., parents, teachers, and early childhood educators) and behaving in ways that support the welfare of others within their peer groups. For example, recent studies have measured toddlers' evaluations of defending actions and of their expectations of the rewards and punishments following an act of aggression (e.g., Geraci, 2020a; Geraci, 2020b; Geraci & Surian, 2021), and have found, largely, that toddlers prefer characters who defend other characters who had themselves been victimized. The results of such studies suggest that even in infancy, babies have a sense of morality, for example, around helping and hurting actions, and highlight a likely innate component to morality, following the natural-tendency view (Bloom, 2013; Hamlin et al., 2007). These findings also highlight the importance of the infancy stage in establishing a foundational understanding of morality in children's lives.

Building from a natural-tendency view of a possible innate aspect of morality, there is also the consideration of how infants do not develop in a vacuum, but rather in a complex ecosystem (Bronfenbrenner, 2005) that includes a lot of early social interaction with parents and caregivers. Following this social-interactional view that early behaviours develop through social interactions (e.g., Dahl, 2015), one of the key ways in which early morality develops is through socialization between caregivers and the child. This socialization begins early in infancy, as infants are tended by parents and caregivers who respond to their needs, such as feeding and comforting (Hammond et al., 2017). One study found that infants with secure caregiver attachment were more likely to look longer at an unresponsive caregiving scenario than those without secure attachment, which suggests the infants with secure attachment were surprised by the unresponsive caregiving scenario (Johnson, Dweck, & Chen, 2007). These findings highlight an association between early caregiver responsiveness, through attachment, and an infant's socio-moral evaluations.

Further, the parenting approaches used in the home serve to shape children's moral behaviour. For example, parents who use warmth and support over power assertion and punishment promote compliant behaviour from an early age (Kochanska & Aksan, 2006). Through childhood, parents' own moral standards such as sensitivity to justice and fairness help to shape the development of moral behaviour according to the child's individual characteristics (e.g, Padilla-Walker & Christensen, 2011). For example, toddlers with a higher fear response internalize attitudes emphasized by parents with gentle discipline techniques, whereas children with lower fear responses do not obtain the same arousal from discipline, and may respond better through discipline that is based on mother–child cooperation and secure attachment (Kochanska, 1993, 1995; Kochanska, Aksan, & Joy, 2007). Early morality, then, while it likely has a natural component, is also constantly being shaped and moulded according to social interactions, for example, responsiveness to need, praise, and withdrawal of attention or affection.

Infants' moral cognition continues to develop as they age and mature (Thompson, 2012), gaining increasingly complex cognitive skills. For example, toddlers' attentive imitation of their mothers during a teaching task, which requires social-cognitive skills such as the theory of mind, is associated with a stronger conscience in early childhood (Forman, Aksan, & Kochanska, 2004). In sum, the development of morality is affected by multiple processes, including the complex interactions with the social environment, including parent and caregiver socialization, as well as cognitive processes and skills.

Early morality and later problematic behaviours? A longstanding question among researchers, parents, and professionals working with young children is what types of early behaviours flag the potential onset for later problematic behaviours? With the identification of early onset behaviours also comes the ability to act in a preventative manner. Research findings with school-age children suggest that moral behaviour during some experimental paradigms is associated with behaviour problems (e.g., Lavoie, Wyman, Crossman, & Talwar, 2018; Mugno et al., 2019). Thus, there is the potential that early morality may act as a predictive indicator for problem behaviours that may become more manifest as the infant or toddler ages. In support of this, studies have found associations between early infant behaviour or emotions, and later problem behaviour. For example, infants' negative reactivity during a mildly distressing activity (Still Face Paradigm, where the mother keeps a still face during an interaction with the child) is associated with fewer symptoms of oppositional-defiance and conduct disorder in childhood (Wagner et al., 2016). More specifically related to morality, infants' preferences for prosocial characters are associated with fewer callous-

unemotional traits in early childhood (Tan, Mikami, & Hamlin, 2018). In our scoping review, we explored whether early measures of morality may have the potential to flag later behaviours of concern through the assessment of studies that included a longitudinal component.

1.3 | Purpose for scoping review of measuring morality in infancy

Given the relation between early trajectories of problematic behaviour and later problematic behaviour through childhood and adolescence (Loeber, 1991; Moffitt & Caspi, 2001), the study of early childhood moral development may be used to inform prevention. One key to being able to conduct this type of study is knowledge of the available methods that can be used to validly, reliably, and feasibly assess moral development. For this reason, we have conducted a scoping review (Colquhoun et al., 2014) of the methods used to study infants' moral development. As mentioned, scoping reviews are a literature gathering method to broadly assess the literature on a specific topic to assist in knowledge synthesis, mapping important concepts, and identifying gaps in the field (Arksey & O'Malley, 2005; Colquhoun et al., 2014).

In our scoping review, we have systematically identified and collected empirical studies that have measured moral development in infancy to generate a thematic overview of the types of methods that have been used as well as the specific constructs within moral development that have been the focus of empirical study. Our secondary aim within this review was to assess whether existing methods in the field may be able to capture variation in moral behaviour that may flag the onset of conduct problems.

2 | METHOD

We conducted a scoping review (Arksey & O'Malley, 2005; Colquhoun et al., 2014) of moral development in infancy and followed the PRISMA-ScR as a guideline. Scoping reviews are useful for broad searches that are not restricted to a fixed question or set of questions (Sucharew & Macaluso, 2019), and that can be used to assess the scope of the research literature (Peters et al., 2015). Given that we wanted to generate a snapshot of the current state of the scope of the methods used to measure moral development and infancy and the specific concepts studied within moral development in infancy, we selected a scoping review approach.

In our searches, we used the specific terms 'moral development' and 'infant' or 'toddler' within the main library catalogue and included databases PsycInfo, Scopus, PsycArticles, and Social Science Citation Index (from 1990 onwards), with 6,748 results returned. We conducted a data text analysis to determine the topical scope of our results returned by the overall search. A simple word frequency analysis of all of the results returned from the title and abstracts demonstrated that common terms yielded from the search included children ($n = 4,334$), social ($n = 2,855$), development ($n = 2,781$), research ($n = 2,236$), health ($n = 2,159$), and infants ($n = 1,578$), which highlights the scope of the returned results.

Screening process. We conducted an initial screen of all of the articles returned by our search criteria above. We retained for further review and inclusion studies that were (a) empirical, (b) had a moral development component (broadly defined to include prosocial: socially accepted and expected behaviour that benefits another person; or antisocial; unaccepted behaviour that can hurt the child or others), (c) had a child sample population of 2 years and under (24 months and under), and (d) that had an English translation accessible. Each article that met this criterion was retained in a citation software management system for coding. This screening process resulted in 70 articles. An additional 19 studies were added from hand searches of key articles that had been returned through the scoping review search, for a final total of 89 articles. All retained articles were assessed for clarity and rigour of the method and results and each article passed a minimum threshold for low risk of bias.

Coding process. Each article was reviewed in duplicate and information regarding the study method and outcome was extracted. We extracted (a) the primary method within the study (experimental: any type of procedure for measuring infants' behaviour in a specific type of paradigm or procedure; observational: structured or unstructured observations of free-activity such as unscripted interactions or play; or questionnaire-based); (b) main outcome type (moral behaviour: the child's own moral behaviour during an activity, e.g., sharing; moral evaluation: the child's moral evaluation of a scenario or event, e.g., whether they determine an action was good or bad; ability to identify emotions: the child identifies emotions; or child's own emotions: the child's own emotionality, e.g., how they reacted to a specific scenario); (c) sample characteristics (participant number and age range); (d) specific aspect of morality measured by the outcome variable (e.g., fairness, empathic concern, and compliance with prohibitions); (e) whether the study was longitudinal; and (f) if longitudinal, whether there was a relation between the time 1 (T1) outcome and the time 2 (T2) outcome to assess for predictive ability and associations between measures of morality and other factors. Although we were most interested in assessing whether early measures of morality could predict later child behaviour, we looked at longitudinal relations more broadly to explore whether there was an evidence basis for predictive associations between morality and other factors (and vice versa, in studies where such was the case). A coding manual is available in Data S2 for further detail.

To synthesize the data, we extracted results from each of the studies in a table format, and subsequently used an evidence mapping approach (Hetrick, Parker, Callahan, & Purcell, 2010; Miake-Lye, Hempel, Shanman, & Shekelle, 2016) to visualize the spread and overlap of aspects of morality that were measured across studies. The organization of themes was guided by theoretical discussions by Thompson (2012) and Malti and Dys (2015), and is discussed further in our results. We used both of these visual outputs to gauge the scope of the current knowledge base on moral development in infancy, as well as identify gaps in the field to generate recommendations for future study.

3 | RESULTS

3.1 | Sample size and age range

There were 185 studies/experiments across 89 articles included in the scoping review. Details about the studies are presented in Table 1. The sample sizes ranged from 4 to 1,204 participants within a study, with a substantial skew towards sample sizes 40 participants and under (see Figure 1 for a breakdown of sample sizes of all experiments, including across time points). The age range of participants in the studies included in the scoping review was 24 months and under. The youngest participants were 1-month old, and the oldest was 2 years old (see Figure 2 for the number of experiments that included participants 1–2 years old vs. those under 1-year old). There were a greater number of experiments that included participants 1–2 years old than experiments with younger infants under 1-year old.

3.2 | Type of study and main outcomes

The majority of the studies included in this scoping review relied on primarily experimental methods ($n = 73$), followed by primarily observational ($n = 14$), and questionnaire-based ($n = 2$). The child's own moral behaviour and the child's moral evaluations were overwhelmingly the predominant primary outcomes within studies (see Figure 3).

Moral behaviour. This approach for measuring moral development relied on infants' actual behaviour, and was captured through a questionnaire ($n = 2$), observational ($n = 10$), and experimental ($n = 28$) methods. A seminal study by Warneken and Tomasello (2006, 2007) employed 6–10 scenarios (or paradigms) with toddlers 14, 18, and 24 months to test toddlers' tendency to behave in helpful ways to the experimenter and found that infants tended

TABLE 1 Study characteristics

	Authors	Year	Type of study	Main outcome	Sample size	Age in months, except when indicated by 'y' (years)	Aspect of moral dev	Longitudinal?	T1 predicts T2?
1	Barragan & Dweck	2014	Experimental	Behaviour	E1: 34 E2: 30[a]	E1: 1–2 y E2: 1–2 y[a]	Altruism of helping others	No	
2	Barrett	2005	Experimental	Own emotions	35	17	Social emotions of guilt and embarrassment	No	
3	Bian et al.	2018	Experimental	Evaluation	E1: 48[a]	E1: 18[a]	Fairness expectations	No	
4	Brooker et al.	2013	Experimental	Behaviour	49	18	Helping behaviour	No	
5	Buon et al.	2014	Experimental	Evaluation	E2: 47[a]	E2: 10[a]	Social evaluation of comforting versus harmful actions	No	
6	Burns & Sommerville	2014	Experimental	Evaluation	E1: 40 E2: 40	E1: 15 E2: 15	Fairness expectations	No	
7	Buyukozer Dawkins et al.	2019	Experimental	Evaluation	E1: 40 E2: 40 E3: 40 E4: 40	E1: 9 E2: 9 E3: 9 E4: 4	Fairness expectations	No	
8	Campbell et al.	2015	Experimental	Own emotions	69	T1: 22[a]	Empathic concern	Yes	Yes
9	Chiarella et al.	2018	Experimental	Behaviour	73	17.1	Helping behaviour	No	
10	Cowell & Decety	2015	Experimental	Evaluation	73	12–24	Social evaluation of prosocial and antisocial agents	No	
11	Dahl	2016	Observational	Behaviour	26	T1: 14 T2: 19 T3: 24	Compliance with maternal prohibitions	Yes	NI
12	Dahl	2015	Observational	Behaviour	E1: 76 E2: 51 E2 T1/T2: 24	E1: 11–25 E2 T1: 13–25 E2 T2: 18–35	Helping	Yes	NI
13	Dahl & Tran	2016	Experimental	Behaviour	E1: 119	E1: 13–15, 18–20, 23–25	Compliance with maternal prohibitions	No	
14	Dahl et al.	2013	Experimental	Evaluation	84	17, 22, 26	Helping prosocial versus antisocial agents	No	

(Continues)

TABLE 1 (Continued)

	Authors	Year	Type of study	Main outcome	Sample size	Age in months, except when indicated by 'y' (years)	Aspect of moral dev	Longitudinal?	T1 predicts T2?
15	DesChamps et al.	2016	Experimental	Evaluation	E1: 36 E2: 22 E3: 36	E1: 15 E2: 15 E3: 13	Social perceptions of fairness	No	
16	Dunfield et al.	2011	Experimental	Behaviour	48	18, 24	Helping, sharing, and comforting	No	
17	Eiden et al.	2002	Experimental	Behaviour	T1: 45 T2: 45 T3: 45	T1: 1 T2: 2 T3: 7	Affect regulation and problem behaviour	Yes	Yes
18	Feldman	2009	Experimental	Behaviour	125	T1: 3 T2: 6 T3: 12 T4: 24[a]	Emotion regulation and self-restraint	Yes	Yes
19	Forman et al.	2004	Experimental	Behaviour	E1 T1: 108 E1 T2: 106[a]	E1 T1: 14 E1 T: 22[a]	Responsive imitation of mother and conscience (internalized conduct and guilt)	Yes	Yes
20	Franchin et al.	2019	Experimental	Evaluation	E1: 65 E2: 66 E3: 24 E4: 23	E1: 20–30 E2: 23–30 E3: 20 E4:20	Understanding of 'good' within helping and fairness	No	
21	Gagné et al.	2018	Observational and experimental	Behaviour	T1: 92[a]	T1: 18[a]	Self-regulation (rule-compatible conduct)	Yes	No
22	Geraci & Surian	2011	Experimental	Evaluation	E1: 16 E2: 20	E1: 10 E2: 16	Sense of fairness in resource allocation	No	
23	Gredebäck et al.	2015	Experimental	Evaluation	E1: 14 E2: 14	E1: 6 E2: 6	Cognitive processes in evaluating pro- and antisocial agents	No	
24	Hamlin	2014	Experimental	Evaluation	E1: 55 E2: 27	E1:3.5–5.5 E2: 3.5–5.5	Social evaluation of helpers versus hinderers	No	
25	Hamlin	2015	Experimental	Evaluation	E1: 48 E2: 50	E1: 6–11 E2: 6–11	Social evaluation of helpers versus hinderers	No	
26	Hamlin	2013	Experimental	Evaluation	E1: 32	E1: 5–8		No	

TABLE 1 (Continued)

	Authors	Year	Type of study	Main outcome	Sample size	Age in months, except when indicated by 'y' (years)	Aspect of moral dev	Longitudinal?	T1 predicts T2?
27	Hamlin & Wynn	2011	Experimental	Evaluation	E2: 32 E3: 32	E2: 5–8 E3: 5–8	Social evaluation of helpers versus hinderers	No	
28	Hamlin et al.	2013	Experimental	Evaluation	90	9–10	Social evaluation of helpers versus hinderers	No	
29	Hamlin et al.	2007	Experimental	Evaluation	E1: 28 E2: 24 E3: 16	E1: 6–10 E2: 6–10 E3: 6–10	Social evaluation of helpers versus hinderers	No	
30	Hamlin et al.	2010	Experimental	Evaluation	E1: 24 E2: 24	E1: 3 E2: 3	Social evaluation of helpers versus hinderers	No	
31	Hart et al.	2004	Experimental	Own emotions	94	4–6	Affect response to jealousy evocation	No	
32	Hatzinikolaou & Murray	2010	Observational	Identify emotions	90	8; 18	Infant sensitivity to maternal emotion shifts	Yes	Yes
33	Heberle et al.	2014	Questionnaire	Behaviour	1,204	12–42	Disruptive behaviours	No	
34	Hecke et al.	2007	Experimental	Behaviour	T1: 52 T2: 52 T3: 52[a]	T1: 12 T2: 15 T3: 24[a]	Social competence and externalizing behaviour	Yes	Yes
35	Hepach et al.	2017	Experimental	Behaviour	E1: 144 E2: 48	E1: 17–19 E2: 17–18	Spontaneous helping	No	
36	Hobbs & Spelke	2015	Experimental	Behaviour	E1: 18 E2: 18 E3: 18 E4: 18 E5: 36	E1: 14 E2: 14 E3: 24 E4: 14 E5: 14	Helping	No	
37	Huang et al.	2017	Experimental	Own emotions	84	10–26	Empathy	No	

(Continues)

TABLE 1 (Continued)

Authors	Year	Type of study	Main outcome	Sample size	Age in months, except when indicated by 'y' (years)	Aspect of moral dev	Longitudinal?	T1 predicts T2?
38 Hyde et al.	2010	Observational and questionnaire	Behaviour	T1: 310[a]	T1: 6–17 / T2: 18 / T3: 24[a]	Empathy, antisocial behaviour, and moral disengagement	Yes	Yes
39 Jin et al.	2018	Experimental	Evaluation	E1: 44 / E2: 36 / E3: 36	E1: 11–13 / E2: 3–5 / E3: 7–9	Social evaluation of responses to need	No	
40 Kanakogi et al.	2017	Experimental	Evaluation	E1: 20 / E2: 20 / E3: 20 / E4: 16 / E5: 16 / E6: 40	E1: 6 / E2: 6 / E3: 6 / E4: 6 / E5: 6 / E6: 6; 10	Evaluation of helping or non-helping agent in an aggressive interaction	No	
41 Kanakogi et al.	2013	Experimental	Evaluation	E1: 40 / E2: 24	E1: 10 / E2: 10	Evaluation of victim, aggressor, or neutral object after an aggressive social interaction	No	
42 Kochanska et al.	2007	Experimental	Behaviour	E1 T1: 106 / E1 T2: 104[a] / E2 T1: 102 / E2 T2: 10[a]	E1 T1: 2 / E1 T2: 33[a] / E2 T1: 7, E2 T2: 15[a]	Fearfulness and the child's moral self	Yes	No
43 Kochanska et al.	2004	Experimental and observational	Own emotions	T1: 108 / T2: 106[a]	T1: 14 / T2: 22[a]	Mother's responsive parenting and child's moral self (conscience, conduct, and cognition)	Yes	Yes
44 Kochanska et al.	2008	Experimental and observational	Behaviour	T1: 102 / T2: 101[a]	T1: 7 / T2: 15[a]	Mother's responsive parenting and child's conscience (rule following and moral judgements) and antisocial behaviour	Yes	Yes
45 Kochanska et al.	1994	Questionnaire and experimental	Behaviour	E1 T1: 62[a]	E1 T1: 21–33[a]	Temperament and conscience	No	
46 Kochanska et al.	2005		Behaviour	T1: 112	T1: 9		Yes	Yes

TABLE 1 (Continued)

Authors	Year	Type of study	Main outcome	Sample size	Age in months, except when indicated by 'y' (years)	Aspect of moral dev	Longitudinal?	T1 predicts T2?
		Observational and experimental		T2: 108 T3: 106[a]	T2: 14 T3: 22[a]	Mother–child responsiveness and child's conscience (moral emotion, conduct, and cognition)		
47 Kochanska et al.	2002	Experimental	Own emotions and behaviour	T1: 108 T2: 106[a]	T1: 18 T2: 22[a]	Guilt and moral self	Yes	Yes
48 Kochanska et al.	2011	Observational and experimental	Own emotions	T1: 101[a]	T1: 15[a]	Genes, maternal responsiveness, and child's moral internalization	Yes	Yes
49 Kochanska et al.	1998	Experimental	Behaviour	108	T1: 8–10 T2: 13–15	Restraint, compliance, and internalization	Yes	Yes
50 Kok et al.	2013	Observational and experimental	Own emotions	T1: 534[a]	T1: 14[a]	Attachment, temperament, and compliance	Yes	Yes
51 Lamb	1993	Observational	Behaviour	T1: 4 T2: 4	T1: 13–14 T2: 21–23	Moral awareness	Yes	NI
52 Lamb & Zakhireh	1997	Observational	Behaviour	45	9–27	Response to distress of other children	No	
53 Licht et al.	2008	Observational	Behaviour	28	T1: 8 T2: 14 T3: 22	Motivations for conflict	Yes	NI
54 Lickenbrock et al.	2013	Experimental	Behaviour	106	T1: 3 T2: 5 T3: 7 T4: 12 T5: 14 T6: 20	Temperament, attachment, and styles of compliance and noncompliance	Yes	Yes
55 Liddle et al.	2015	Experimental	Behaviour	25	7–10	Response to peer distress	No	
56 Lucca et al.	2018	Experimental	Behaviour	E1: 14 E2: 16	E1: 13 E2: 17	Fairness expectations	No	

(Continues)

TABLE 1 (Continued)

	Authors	Year	Type of study	Main outcome	Sample size	Age in months, except when indicated by 'y' (years)	Aspect of moral dev	Longitudinal?	T1 predicts T2?
57	Lyubchik & Schlosser	2010	Experimental	Identify emotions	54	16	Empathy	No	
58	Meristo & Surian	2014	Experimental	Evaluation	E1: 16 E2: 16 E3: 16	E1: 9–10 E2: 9–10 E3: 9–10	Evaluation of equal/unequal or fair/ unfair distribution of goods	No	
59	Meristo & Surian	2013	Experimental	Evaluation	E1: 16 E2: 16 E3: 32	E1: 10 E2: 10 E3: 10	Evaluation of third party behaviour towards fair/unfair behaviour	No	
60	Meristo et al.	2016	Experimental	Evaluation	E1: 16 E2: 16 E3: 16 E4: 16	E1: 10 E2: 10 E3: 10 E4: 10	Evaluation of equal/unequal distribution of goods	No	
61	Morales et al.	2005	Experimental	Behaviour	22	T1: 6 T2: 24	Joint attention and emotion regulation	Yes	Yes
62	Morrell & Murray	2003	Observational and experimental	Behaviour	59	T1: 2 T2: 4 T3: 9[a]	Emotional dysregulation and conduct problems	Yes	Yes
63	Nichols et al.	2010	Experimental	Behaviour	T1: 23 T2: 24 T3: 21	T1: 12 T2: 18 T3: 24	Response to toys based on peer's emotional affect	No	
64	Nichols et al.	2015	Experimental	Behaviour	71	12; 18; 24	Response to other children in distress	No	
65	Paquette-smith & Johnson	2016	Experimental	Evaluation	E1: 20 E2: 20 E3:16	10–11	Preference for prosocial tone	No	
66	Paulus et al.	2015	Experimental	Behaviour	T1: 65 T2: 61 T3: 62[a]	T1: 7 T2: 18 T3: 24[a]	Distress understanding and sharing	Yes	Yes
67	Pérez-Edgar et al.	2017	Experimental	Behaviour	145	4–24	Temperament and emotional face gaze fixation	No	

TABLE 1 (Continued)

|	Authors	Year	Type of study	Main outcome	Sample size	Age in months, except when indicated by 'y' (years)	Aspect of moral dev	Longitudinal?	T1 predicts T2?
68	Poehlmann et al.	2012	Observational	Behaviour	109	T1: 9[a]	Toddler compliance and behaviour problems	Yes	Yes
69	Salvadori et al.	2015	Experimental	Evaluation	E1: 24 E2: 24	E1: 8–9 E2: 8–9	Evaluation of helping or hindering	No	
70	Schmidt & Sommerville	2011	Experimental	Evaluation	37	15	Sensitivity to fairness (resource allocation)	No	
71	Scola et al.	2015	Experimental	Evaluation	39	13–23[a]	Preference for prosocial versus antisocial agent	No	
72	Shimizu et al.	2018	Experimental	Evaluation	159	6; 9; 12; 15–18	Evaluation of helpers and hinderers	No	
73	Sloane et al.	2012	Experimental	Behaviour	E1: 48 E2: 54	E1: 19 E2: 21	Expectations of fairness (resource allocation)	No	
74	Sodian et al.	2016	Experimental	Evaluation	T1: 77 T2: 77 T3: 77[a]	T1: 7 T2: 18 T3: 24[a]	Intentional agency and moral intentions	Yes	Yes
75	Sommerville et al.	2013	Experimental	Evaluation	E1: 48 E2: 34	E1: 12 E2: 15	Expectations of fairness (resource allocation)	No	
76	Spinrad & Stifter	2006	Observational and experimental	Own emotions and behaviour	T1: 98 T2: 98	T1: 10 T2: 18	Negative emotionality and empathy	Yes	Yes
77	Stavans & Baillargeon	2019	Experimental	Evaluation	E1: 48 E2: 48 E3: 24	E1: 16–18 E2: 16–18 E3: 16–18	Fairness expectations of leaders	No	
78	Steckler et al.	2017	Experimental	Evaluation	E1: 32 E2: 16 E3: 64 E4: 64	8–9	Social evaluations of repeat prosocial versus antisocial agents	No	
79	Surian & Franchin	2017	Experimental	Evaluation	64	13–17; 18–22	Fairness (equal distributions) to prosocial versus antisocial agents	No	
80	Surian et al.	2018	Experimental	Evaluation	32	15	Fairness (equal distributions) to prosocial versus antisocial agents	No	

(Continues)

TABLE 1 (Continued)

|	Authors	Year	Type of study	Main outcome	Sample size	Age in months, except when indicated by 'y' (years)	Aspect of moral dev	Longitudinal?	T1 predicts T2?
81	Tan et al.	2018	Experimental	Evaluation	T1: 63[a]	T1: 12[a]	Evaluation of fairness (equal or unequal distribution) of helping versus hindering agents; Social evaluation of helping versus hindering agent and social-moral competence	Yes	Yes
82	Vaish et al.	2009	Experimental	Evaluation and behaviour	64	17–18; 23–26	Social evaluation of helping versus hindering and helping behaviour	No	
83	Van de Vondervoort et al.	2018	Experimental	Behaviour and own emotions	E1: 64 E2: 64	E1: 19–20 E2: 19–20	Giving behaviour and own emotions after giving/withholding	No	
84	Wagner et al.	2016	Experimental and observational	Behaviour	185	T1: 6 T2: 24[a]	Mutual gaze orientations, stress responsivity, and OD/CU behaviours	Yes	Yes
85	Wang & Henderson	2018	Experimental	Evaluation	E1: 36 E2A: 36 E2B: 12	E1: 16–18 E2A: 16–18 E2B: 16–18	Fairness expectations	No	
86	Warneken & Tomasello	2006	Experimental	Behaviour	24	18	Altruism (helping behaviour)	No	
87	Warneken & Tomasello	2007	Experimental	Behaviour	E1: 24 E2: 24	E1: 14 E2: 14	Helping and cooperation	No	
88	Zahn-Waxler et al.	1992	Experimental	Behaviour	T1: 30 T2: 30 T3: 30	T1: 13–15 T2: 18–20 T3: 23–25	Expressions of concern and prosocial behaviours (helping, sharing, and comforting)	Yes	Yes
89	Ziv & Sommerville	2017	Experimental	Evaluation	E1: 32[a] E3: 34	E1: 12[a] E3: 9	Fairness (resource allocation)	No	

Note: E1: Experiment 1; E2: Experiment 2, and so forth. T1: Time 1 (for longitudinal studies); T2: Time 2, and so forth. NI: not statistically indicated for various reasons (e.g., sample too small, qualitative study).

Abbreviation: OD/CU, Oppositional defiant/callous-unemotional.

[a]For additional details, including time points that were outside of the 2 years and under (24 months and under) sample inclusion criteria, please consult paper.
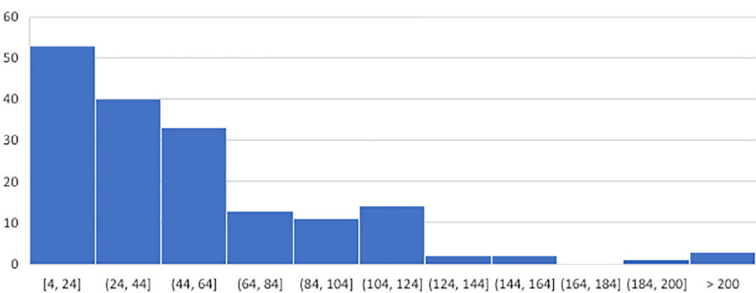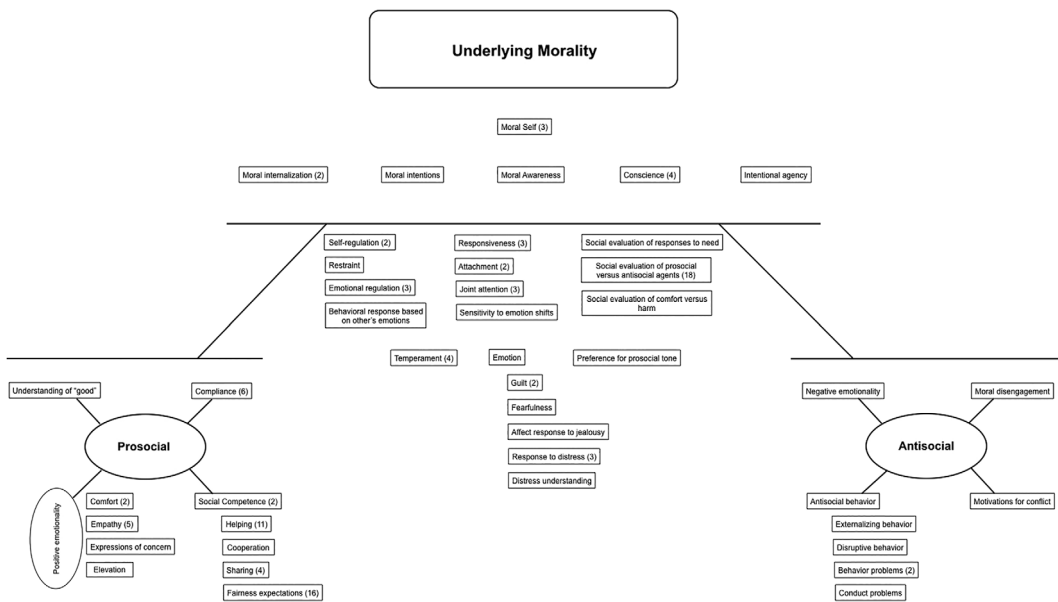
FIGURE 1 Sample size distribution (number of participants) across experiments listed in Table 1
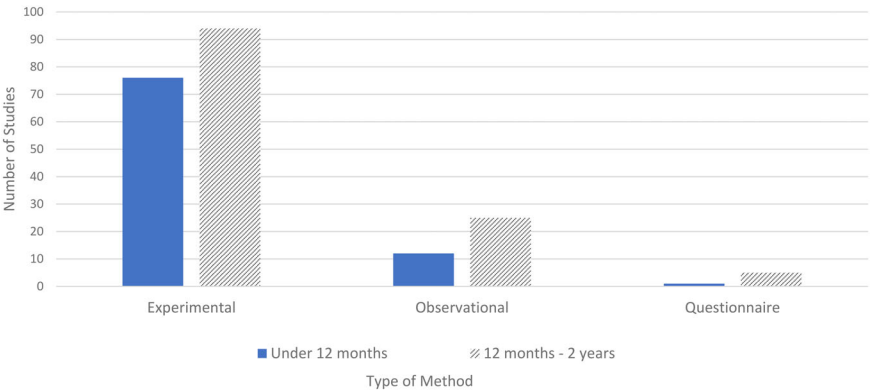
FIGURE 2 Distribution of studies with infants 1–2 years and infants under 1 year, according to the type of method
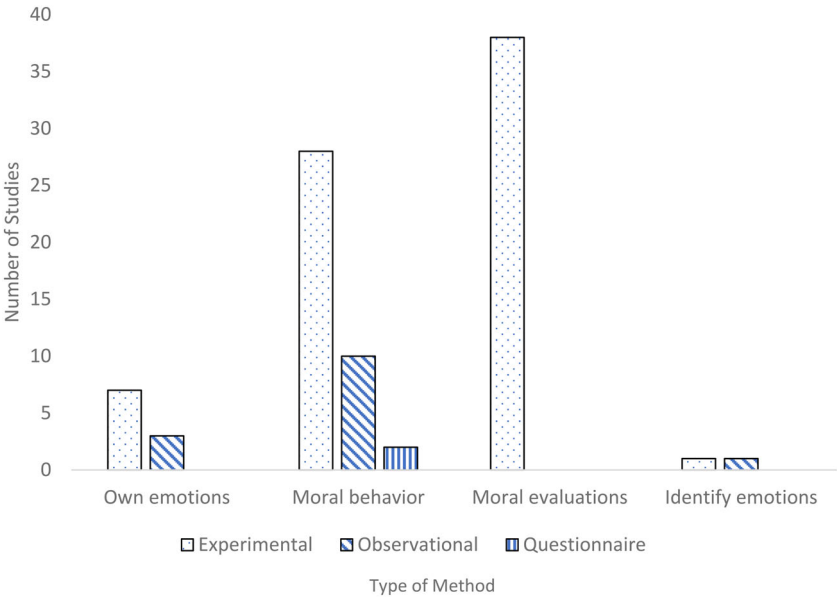
FIGURE 3 Number of studies according to the type of method and main measurement outcome

FIGURE 4 Concept map of specific aspects of moral development studied. Concepts are organized thematically, with main concepts in large bold text, and related concepts in close proximity

to help the experimenter with out-of-reach objects, but not with more complex tasks. Other studies have measured behaviours such as cooperation (Warneken & Tomasello, 2007), compliance with instructions (e.g., Dahl, 2016; Lickenbrock et al., 2013; Poehlmann et al., 2012), comforting another (e.g., Chiarella & Poulin-Dubois, 2018; Dunfield, Kuhlmeier, O'Connell, & Kelley, 2011), and sharing (e.g., Dunfield et al., 2011; Paulus et al., 2015; Ziv & Sommerville, 2017). Table 1 presents a complete list of the aspect of moral development that was the focus of study within each article. Figure 4 maps the moral development concepts and includes a frequency count for overlapping outcomes.

There was considerable variance in the methodologies employed to study infants' moral behaviour. Some studies opted to use an experimental approach through the development of standardized paradigms to assess infants' responses during a controlled laboratory scenario, such as dropping an item and watching to see if the child would assist (e.g., Warneken & Tomasello, 2006). Other studies used an observational approach and observed how infants would react and respond to naturalistic events that emerged during the assessment, such as complying with parental requests or instructions (e.g., Dahl, 2016; Poehlmann et al., 2012) to measure moral behaviour within the natural context of daily life.

Moral evaluations. This approach for measuring moral development captured infants' evaluations of moral situations (broadly defined again to include prosocial or antisocial situations) and was entirely captured using experimental methods ($n = 38$). The predominant method was through evaluating infants' preferences for prosocial (vs. neutral or antisocial) characters by measuring indicators such as looking time and reaching for a prosocial character (over an antisocial character), which was a common measurement method with adaptations. A seminal study in the field by Hamlin et al. (2007) found that 6- and 10-month old infants were more likely to reach for a character who had behaved prosocially (helped another character), over a neutral or negative character (who hindered another character from reaching their goal). They also found that infants were more likely to look longer at an event in which the main character approached the negative character (who had hindered it previously), which following a VOE hypothesis likely indicates the infants were surprised by the event. The authors use these results to argue for innate social evaluation preferences in infants (as a 'foundation to moral thought and action' p. 557).

Another common aspect of moral development measured using evaluative methods was infants' fairness expectations, specifically of the fair allocation of resources. Following a similar methodology to the experimental paradigms employed for studying infants' preferences for prosocial versus antisocial agents, of looking time and reaching for

characters, these studies found that infants tend to prefer agents who allocate resources similarly across group members (e.g., Burns & Sommerville, 2014; Geraci & Surian, 2011; Meristo et al., 2016; Meristo & Surian, 2014).

Infants' moral emotions. Two studies focused on infants' ability to identify emotions (*n* = 2) and 10 studies focused on infants' own emotions (*n* = 10). Studies looked at infants' abilities to detect maternal shifts in emotion (Hatzinikolaou & Murray, 2010) and at their ability to identify empathy (Lyubchik & Schlosser, 2010), using observational methods and experimental methods, respectively. Within studies that measured infants' own emotions, guilt (Barrett, 2005; Kochanska et al., 2002) and empathy (Campbell et al., 2015; Huang et al., 2017) were the two most common emotions studied, with two studies on each, all employing an experimental methodology to measure the infants' own emotions within a controlled setting.

Evidence mapping of moral development outcomes. We used an evidence-mapping methodology (Hetrick et al., 2010; Miake-Lye et al., 2016) to visually organize and synthesize the scope of moral development that was captured by the studies in this review, as well as to identify gaps (discussed in the next section). Figure 4 presents our evidence map of the specific measures within the scope of moral development that were studied. Main concepts are represented by large bold text and are organized thematically. The organization of themes was guided by theoretical discussions by Thompson (2012) and Malti and Dys (2015).

We conceptually organized first the centre, or core, as an underlying sense of morality, that is, the place from which moral decisions or behaviours flow. Concepts above the top line represent aspects of morality that are fundamental, but not explicitly visible per se. In other words, they are concepts that need to be operationalized to be measured. This includes topics such as the moral self (three articles in the review), the conscience (four articles in the review), and moral awareness.

Directly below 'underlying morality', we placed aspects of moral development that can be considered operationalizations of the core of morality and that are not necessarily 'good' nor 'bad', but rather operationalized concepts that stem from a moral core. This includes self-regulation (two articles in the review), social evaluations (20 articles across types of social evaluations), and moral emotions that stem from moral awareness, such as guilt (two articles), and emotional responses to jealousy (one article).

To the left and right (respectively), and below these concepts, we have placed prosocial components and antisocial components to visually depict our conceptualization of them stemming from a deeper underlying sense of morality and awareness, and to distinguish them from moral awareness. Prosocial components include behaviours such as compliance (six articles), helping (11 articles), and fairness expectations (13 articles). Antisocial components include behaviours under the umbrella of antisocial behaviour (e.g., externalizing behaviour and disruptive behaviour) and moral disengagement.

Longitudinal findings. There were 30 studies included in our review that reported a longitudinal component. Studies commonly used longitudinal methods to examine moral development as an outcome at a later testing point with early predictors such as parenting (Kochanska, Forman, Aksan, & Dunbar, 2005; Kochanska, Kim, Barry, & Philibert, 2011), attachment (Kok et al., 2013), and emotion regulation (Eiden, Lewis, Croff, & Young, 2002; Feldman, 2009; Morrell & Murray, 2003; Spinrad & Sifter, 2006). No study repeated the same task and measure longitudinally (or an adapted version that might avoid infant familiarity with the task) to track a specific aspect of moral development.

Eight longitudinal studies had a specific focus on behaviour problems (Eiden et al., 2002; Hecke et al., 2007; Hyde et al., 2010; Kochanska et al., 2008; Morrell & Murray, 2003; Poehlmann, 2012; Tan et al., 2018; Wagner et al., 2016). These studies help to shed light on whether early morality may be useful and have a predictive ability to identify later problematic behaviours. For example, Tan et al. (2018) found that infants' early identification of prosocial characters was associated with fewer callous-unemotional traits in early childhood. Of these studies, six were experimental in focus, five were observational in focus, and one was questionnaire-based. Four studies were primarily multi-method, combining these methods to measure morality. Together, these findings suggest that there are direct (e.g., attachment and emotion regulation) and indirect (e.g., parenting) factors related to moral development that can be used longitudinally to predict moral, pro-, and antisocial behaviour, but that studies have not tracked a specific element of moral development over time, whether with the same task or adapted ones.

4 | DISCUSSION

4.1 | Summary

The purpose of this scoping review was to survey the field for studies of moral development in infancy and to synthesize the types of methods and specific aspects of moral development within this review. We found that the overwhelming majority of studies used an experimental methodology to collect information on moral development, and within this, infants' actual behaviour and their evaluations were the most common sources of information. An evidence map depicts concept delineation between studies and presents concepts as related to an underlying moral sense, as prosocial components (emotion and behaviour), and as antisocial components (emotion and behaviour). Just under one-third of the studies were longitudinal, and a high percentage of them reported a statistically significant relation between moral development measurements at different time points.

4.2 | Measures of moral development

Infants' own behaviour. We found that a common method of studying moral development was studying infants' own behaviour. Within these studies, the focus tended to lean towards the measurement of prosocial behaviours in infancy, for example, sharing with another individual and comforting another individual (Dunfield et al., 2011; Paulus et al., 2015; Ziv & Sommerville, 2017). The methods used for measuring moral behaviour varied between unstructured observations of the infant's behaviour, questionnaires completed by parents about the infant's behaviour, to structured experimental paradigms to test the infant's behavioural response to a specific scenario. Of note, one of the great challenges in measuring infants' moral behaviour is that many developmental skills are needed to truly assess moral behaviour. For example, regarding sharing specifically, researchers have found that young children (∼2–5 years old) are not able to share fairly, not out of a desire to *not* share, but because their numerical abilities are still developing (i.e., not being fully competent in complex mathematics of equitable resource distribution; Chernyak, Harris, & Cordes, 2018). This highlights some of the methodological challenges inherent to studying moral behaviour with infants.

One positive finding to emerge regarding infants' moral behaviour is that it is highly feasible to measure moral behaviour, even in infancy. Studies included in this review measured an aspect of moral behaviour even as early as within the first 4 months of life (e.g., Eiden et al., 2002; Morrell & Murray, 2003), which opens the possibility to developing future measures to capture infants' early moral behaviour, for example, empathy, resource allocation, and compliance, in novel paradigms.

Infants' moral evaluations. The predominant method of measuring moral evaluations was of studying infants' social evaluations of a helping agent versus a hindering agent (Hamlin et al., 2007), which although not directly a moral question per se (Carpendale & Hammond, 2016), does stem from a moral awareness of right versus wrong social actions. Since this study, Holvoet, Scola, Arciszewski, and Picard (2016) sourced an additional 27 experiments across 12 studies (publications) that employed this or a similar methodology with mixed results (for a synopsis of the additional studies see Holvoet et al., 2016 for a systematic review). Specifically, Holvoet and colleagues report that, of the 27 experiments, 18 found a preference for a prosocial character. The inconsistent findings may be highlighting another underlying methodological challenge, which is that it is difficult to capture infants' evaluations given their developmental capacity in aspects such as attention span, emotional self-regulation for any uncomfortable distractions (fatigue and hunger), and pre-verbal ability.

Perhaps because of these inherent measurement difficulties, the series of findings from the helper/hinderer paradigm have been critiqued methodologically and conceptually. For example, there is debate over the interpretation of looking time, questions of whether results are driven by perceptual reasons such as whether the main climber's gaze is looking towards the target goal, rather than social evaluative reasons (Holvoet et al., 2016). Despite the

inconsistent findings, a large number of studies have successfully (in terms of design, low attrition, completion rates, and theoretically supported interpretable findings) employed an evaluative method, which suggests that it is possible to measure infants' social evaluations, but that there is still room for improvement in the specifics of measurement methods. For example, it may be of interest for studies to employ alternative methods for measuring social evaluations that do not rely on elements such as looking time or vignette-based scenarios to develop a broader evidence based on very early morality that can then be judged against the findings from the helper/hinderer paradigm.

Some studies in the field have integrated measures such as eye-tracking and pupil dilation as a measure of arousal (e.g., Cowell & Decety, 2015; Geraci & Surian, 2011; Gredeback et al., 2015; Kanakogi et al., 2017), and future studies may consider comparing results from a social evaluative task using the standard measure of looking time, in comparison with neurocognitive measures (e.g., eye-tracking and functional near-infrared spectroscopy; fNIRs), as well as observational methods such as providing doll figures of the vignette characters for the infant to play with and measuring the infant's behaviour. The considerable interest in measuring infant moral evaluations, gauged by the number of studies in this area, suggests that the resolution of the methodological challenges in this area should be a key focus of future research. Future studies will be able to continue building and developing on the ground-breaking work that has already begun in this area, for the benefit of further in-depth understanding of how infants think and evaluate moral situations.

Infants' moral emotions. Within studies that measured moral emotions, experimental methods for measuring infants' own emotions were most commonly employed. Across the 12 studies included that studied moral emotions (including studies that measured infants' own emotions and infants' ability to identify emotions), empathy was the most highly studied moral emotion (three studies: Campbell et al., 2015; Huang et al., 2017; Lyubchik & Schlosser, 2010), followed by guilt (two studies: Barrett, 2005; Kochanska et al., 2002). Experimental methods were commonly used, but no specific paradigm or scenario predominated. The study of infants' moral emotions may be a promising avenue for future research on infants' morality given the relatively restricted range of emotions that have been studied, as well as the fact that moral emotions can be measured without the need for language skills, which makes it possible to study even in very young infants.

4.3 | Longitudinal measures

Within the scoping review, just over a third of the studies included a longitudinal component. Given that moral behaviour during experimental paradigms differentiates between children with typical versus atypical levels of problem behaviours (e.g., Lavoie et al., 2018; Mugno et al., 2019), one of the secondary aims of our review was to explore the possibility that a measure of moral development may be able to act as an early flag to indicate a potential onset of behaviour problems that could be caught early enough for preventative support and action. We found that it was possible to predict moral behaviour and awareness (e.g., conscience), specifically problem behaviours, using factors directly (e.g., attachment and emotion regulation), and indirectly (e.g., parenting) related to morality. The majority of such studies focused on measuring infants' behaviour (e.g., Eiden et al., 2002; Forman et al., 2004; Kochanska, Aksan, Knaack, & Rhines, 2004), but two studies did measure infants' evaluations (Sodian et al., 2016; Tan et al., 2018). However, no study focused on using a measure of morality to predict morality at a different time point. This is a key direction for future studies that include a longitudinal component and that also focus on the onset and development of behaviour problems.

4.4 | Limitations and future directions

A scoping review is by nature very broad, and while it is a strength to have an overview of a particular field, there are also limitations to trying to gather all of the information. For example, given that aspects of moral development

do not map directly onto the literal term 'moral', it is possible that some studies of relevance were not captured by our search strategy. However, we did find that studies reporting prosocial and antisocial behaviour components were returned in our search results, which does help to minimize the number of articles that may have been missed. Future searches could focus on more specific aspects of moral development, such as virtues (e.g., courage and gratitude) to provide more in-depth analyses of infants' moral development as well as the factors that may predict moral development.

We also looked to gather and organize the information thematically, but in doing so, some of the more nuanced interrelations were not given much emphasis or focus. This scoping review would be well complemented by more directed searches particularly on the longitudinal aspect of moral development and the information that this can provide to help improve behavioural outcomes for children over the long term.

Further, it is possible that 'morality' in infancy looks somewhat different than morality in other age groups given preverbal abilities and other developmentally contingent challenges. That is, infants cannot yet tell the researchers at this age what they think morality or moral behaviour is, but rather assumptions are made based on observations of infants' behaviour in naturalistic contexts or in experimental paradigms that have to be carefully designed to exclude alternative explanations. These limitations need to be carefully considered in the design of future research methods and in the interpretation of findings.

Based on the results of the studies included in this scoping review, there are several promising areas for future research. One area that stands out as a key gap is a need for studies that measure both infants' evaluation of a moral subject and their behaviour. Given that an important aspect of moral development is both knowing and doing—I know what is moral and my behaviour will either reflect or challenge that knowing—there is a gap to be filled in infancy studies that consider both the *evaluation* of moral subjects and then the *doing* of moral tasks. In our scoping review, one study included both an evaluation and behavioural measure (Vaish, Carpenter, & Tomasello, 2009), and there is more space in the literature to develop this harmonic or discordant relation further.

Another gap that stands out is topical. The field of moral development has a large scope with many relevant topics to study and to understand how they develop from infancy through childhood. Roughly one-quarter of the studies included some component of emotionality and covered quite a broad spectrum within this, which limits the current state of understanding of infants' moral emotions as well as the precursors to such emotions. Further, the development of specific virtues such as courage and gratitude, and also of broader concepts such as justice has ample room for future studies on the development of morality in infancy. Additional research coverage and evidence in these areas will provide valuable information that will increase a deeper understanding of how infants develop a moral sense and of how this development may be able to point as an indicator to early trajectories of behavioural problems before they become a concern.

In terms of methodological gaps, there is a great amount of space for the development of longitudinal studies to track moral development over time, beginning in infancy. Although roughly one-third of the studies included in this scoping review did include a longitudinal element, no studies examined moral development longitudinally at multiple time points. This would be of considerable value for future studies to consider as it would provide information about emerging morality as well as how morality develops and changes over time, including the types of factors that influence how morality changes from early infancy through childhood.

5 | CONCLUSIONS

The purpose of our scoping review was to gather information about methods for studying moral development and morality in infancy, and to synthesize and present this information for the benefit of assisting the design of future studies, particularly longitudinal studies, on the design considerations for measuring morality in infancy. We found that measurement of infants' moral evaluations and behaviour are quite feasible, and that both of these aspects may

also be measured longitudinally with the potential to act as an early flag indicator for future problem behaviours, which is an area for further study.

AUTHOR CONTRIBUTIONS

Jennifer Lavoie: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing – original draft; writing – review and editing. **Aja L Murray:** Conceptualization; data curation; funding acquisition; methodology; project administration; resources; supervision; writing – original draft; writing – review and editing. **Guy Skinner:** Data curation; methodology; writing – original draft. **Emilia Janiczek:** Data curation; methodology; writing – original draft.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Jennifer Lavoie https://orcid.org/0000-0002-2247-7905

REFERENCES

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*, 19–31.

Bloom, P. (2013). *Just babies: The origins of good and evil*. New York, NY: Crown.

Bronfenbrenner, U. (2005). Ecological systems theory (1992). In U. Bronfenbrenner (Ed.), *Making human beings human: Bioecological perspectives on human development* (pp. 106–173). Thousand Oaks; London: Sage Publications Ltd.

Buon, M., Seara-Cardoso, A., & Viding, E. (2016). Why (and how) should we study the interplay between emotional arousal, theory of mind, and inhibitory control to understand moral cognition? *Psychonomic Bulletin & Review*, *23*, 1660–1680. https://doi.org/10.3758/s13423-016-1042-5

Burns, M. P., & Sommerville, J. (2014). "I pick you": The impact of fairness and race on infants' selection of social partners. *Frontiers in Psychology*, *5*, 93. https://doi.org/10.3389/fpsyg.2014.00093

Campbell, S. B. (1991). Longitudinal studies of active and aggressive preschoolers: Individual differences in early behavior and outcome. In D. Cicchetti & S. L. Toth (Eds.), *Rochester symposium on developmental psychopathology* (pp. 57–90). Hillsdale, NJ: Erlbaum.

Carpendale, J., & Hammond, S. I. (2016). The development of moral sense and moral thinking. *Current Opinion in Pediatrics*, *28*, 743–747. https://doi.org/10.1097/MOP.0000000000000412

Chernyak, N., Harris, P. L., & Cordes, S. (2018). Explaining early moral hypocrisy: Numerical cognition promotes equal sharing behavior in preschool-aged children. *Developmental Science*, *22*, e12695. https://doi.org/10.1111/desc.12695

Chiarella, S. S., & Poulin-Dubois, D. (2018). "Are you really sad?" infants show selectivity in their behaviors toward an unconventional emoter. *Infancy*, *23*, 453–470. https://doi.org/10.1111/infa.12230

Colquhoun, H. L., Levac, D., O'Brien, K. K., Straus, S., Tricco, A. C., Perrier, L., ... Moher, D. (2014). Scoping reviews: Time for clarity in definition, methods, and reporting. *Journal of Clinical Epidemiology*, *67*, 1291–1294.

Dahl, A. (2015). The developing social context of infant helping in two U.S. samples. *Child Development*, *86*, 1080–1093. https://doi.org/10.1111/cdev.12361

Dahl, A. (2016). Mothers' insistence when prohibiting infants from harming others in everyday interactions. *Frontiers in Psychology*, *7*, 1448. https://doi.org/10.3389/fpsyg.2016.01448

Dahl, A., & Killen, M. (2018). A developmental perspective on the origins of morality in infancy and early childhood. *Frontiers in Psychology*, *9*, 1–6. https://doi.org/10.3389/fpsyg.2018.01736

Dunfield, K., Kuhlmeier, V. A., O'Connell, L., & Kelley, E. (2011). Examining the diversity of prosocial behavior: Helping, sharing, and comforting in infancy. *Infancy*, *16*, 227–247. https://doi.org/10.1111/j.1532-7078.2010.00041.x

Dunfield, K. A., & Kuhlmeier, V. A. (2010). Intention-mediated selective helping in infancy. *Psychological Science*, *21*, 523–527. https://doi.org/10.1177/0956797610364119

Eiden, R. D., Lewis, A., Croff, S., & Young, E. (2002). Maternal cocaine use and infant behavior. *Infancy*, 3, 77–96. https://doi.org/10.1207/S15327078IN0301_4

Forman, D. R., Aksan, N., & Kochanska, G. (2004). Toddlers' responsive imitation predicts preschool-age conscience. *Psychological Science*, 15, 699–704. https://doi.org/10.1111/j.0956-7976.2004.00743.x

Geraci, A. (2020a). How do toddlers evaluate defensive actions toward third parties? *Infancy*, 25, 910–926. https://doi.org/10.1111/infa.12367

Geraci, A. (2020b). How evaluation of protective third-party interventions and the relationship context interact at 21 months. *European Journal of Developmental Psychology*, 17, 556–577. https://doi.org/10.1080/17405629.2019.1680360

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, 14, 1012–1020. https://doi.org/10.1111/j.1467-7687.2011.01048.x

Geraci A., & Surian, L. (2021). Toddlers' expectations of third-party punishments and rewards following an act of aggression. *Aggressive Behavior*, 47, 521–529. https://doi.org/10.1002/ab.21979

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26, 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Greene, J. D. (2007). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 3: The neuroscience of morality: Emotion, disease, and development*. Cambridge, MA: MIT Press.

Hamlin, J. K. (2013). Moral judgment and action in preverbal infants and toddlers: Evidence for an innate moral core. *Current Directions in Psychological Science*, 22, 186–193. https://doi.org/10.1177/0963721412470687

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation in preverbal infants. *Nature*, 450, 557–559. https://doi.org/10.1038/nature06288

Hammond, S. I., Al-Jbouri, E., Edwards, V., & Feltham, L. E. (2017). Infant helping in the first year of life: Parents' recollection of infants' earliest prosocial behaviors. *Infant Behavior and Development*, 47, 54–57. https://doi.org/10.1016/j.infbeh.2017.02.004

He, Z., Bolz, M., & Baillargeon, R. (2011). False-belief understanding in 2.5-year-olds: Evidence from violation-of-expectation change-of-location and unexpected-contents tasks: False-belief understanding in 2.5-year-olds. *Developmental Science*, 14, 292–305. https://doi.org/10.1111/j.1467-7687.2010.00980.x

Hepach, R., Vaish, A., & Tomasello, M. (2017). Children's intrinsic motivation to provide help themselves after accidentally harming others. *Child Development*, 88, 1251–1264. https://doi.org/10.1111/cdev.12646

Hespos, S. J., & Baillargeon, R. (2008). Young infants' actions reveal their developing knowledge of support variables: Converging evidence for violation-of-expectation findings. *Cognition*, 107, 304–316. https://doi.org/10.1016/j.cognition.2007.07.009

Hetrick, S. E., Parker, A. G., Callahan, P., & Purcell, R. (2010). Evidence mapping: Illustrating an emerging methodology to improve evidence-based practice in youth mental health. *Journal of Evaluation in Clinical Practice*, 16, 1025–1030. https://doi.org/10.1111/j.1365-2753.2008.01112.x

Holvoet, C., Scola, C., Arciszewski, T., & Picard, D. (2016). Infants' preference for prosocial behaviors: A literature review. *Infant Behavior & Development*, 45, 125–139. https://doi.org/10.1016/j.infbeh.2016.10.008

Jin, K., & Baillargeon, R. (2017). Infants possess an abstract expectation of ingroup support. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 8199–8204. https://doi.org/10.1073/pnas.1706286114

Jin, K., Houston, J. L., Baillargeon, R., Groh, A. M., & Roisman, G. I. (2018). Young infants expect an unfamiliar adult to comfort a crying baby: Evidence from a standard violation-of-expectation task and a novel infant-triggered-video task. *Cognitive Psychology*, 102, 1–20. https://doi.org/10.1016/j.cogpsych.2017.12.004

Johnson, S. C., Dweck, C. S., & Chen, F. S. (2007). Evidence for infants' internal working models of attachment. *Psychological Science*, 18, 501–502. https://doi.org/10.1111/j.1467-9280.2007.01929.x

Killen, M., & Rutland, A. (2011). Understanding children's worlds. In *Children and social exclusion: Morality, prejudice, and group identity*. West Sussex: Wiley Blackwell. https://doi.org/10.1002/9781444396317

Killen, M., & Smetana, J. G. (2015). Origins and development of morality. In M. E. Lamb & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Socioemotional processes* (pp. 701–749). Hoboken, NJ: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118963418.childpsy317

Kochanska, G. (1993). Toward a synthesis of parental socialization and child temperament in early development of conscience. *Child Development*, 64, 325–347. https://doi.org/10.2307/1131254

Kochanska, G. (1995). Children's temperament, mothers' discipline, and security of attachment: Multiple pathways to emerging internalization. *Child Development*, 66, 597–615. https://doi.org/10.2307/1131937

Kochanska, G., & Aksan, N. (2006). Children's conscience and self-regulation. *Journal of Personality*, 74, 1587–1618. https://doi.org/10.1111/j.1467-6494.2006.00421.x

Kochanska, G., Aksan, N., & Joy, M. E. (2007). Children's fearfulness as a moderator of parenting in early socialization: Two longitudinal studies. *Developmental Psychology*, 43, 222–237. https://doi.org/10.1037/0012-1649.43.1.222

Kochanska, G., Aksan, N., Knaack, A., & Rhines, H. M. (2004). Maternal parenting and children's conscience: Early security as moderator. *Child Development*, 75, 1229–1242. https://doi.org/10.1111/j.1467-8624.2004.00735.x

Kochanska, G., Forman, D. R., Aksan, N., & Dunbar, S. B. (2005). Pathways to conscience: Early mother–child mutually responsive orientation and children's moral emotion, conduct, and cognition. *Journal of Child Psychology and Psychiatry*, 46, 19–34. https://doi.org/10.1111/j.1469-7610.2004.00348.x

Kochanska, G., Kim, S., Barry, R. A., & Philibert, R. A. (2011). Children's genotypes interact with maternal responsive care in predicting children's competence: Diathesis-stress or differential susceptibility? *Development and Psychopathology*, 23, 605–616. https://doi.org/10.1017/S0954579411000071

Kok, R., van IJzendoorn, M. H., Linting, M., Bakermans-Kranenburg, M. J., Tharner, A., Luijk, M. P. C. M., ... Tiemeier, H. (2013). Attachment insecurity predicts child active resistance to parental requests in a compliance task. *Child: Care, Health and Development*, 39, 277–287. https://doi.org/10.1111/j.1365-2214.2012.01374.x

Krettenauer, T. (2012). Linking moral emotion attributions with behavior: Why "(un)happy victimizers" and "(un)happy moralists" act the way they feel. *New Directions for Youth Development*, 2012, 59–74. https://doi.org/10.1002/yd.20039

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh*. New York, NY: Better Books.

Lapsley, D., & Carlo, G. (2014). Moral development at the crossroads: New trends and possible futures. *Developmental Psychology*, 50, 1–7. https://doi.org/10.1037/a0035225

Lavoie, J., Leduc, K., Crossman, A., & Talwar, V. (2016). Do as I say and not as I think: Parent socialisation of lie-telling behavior. *Children & Society*, 30, 253–264. https://doi.org/10.1111/chso.12139

Lavoie, J., Nagar, P. M., & Talwar, V. (2017). From Kantian to Machiavellian deceivers: Development of children's reasoning and self-reported use of secrets and lies. *Childhood*, 24, 197–211. https://doi.org/10.1177/0907568216671179

Lavoie, J., Wyman, J., Crossman, A. M., & Talwar, V. (2018). Lie-telling as a mode of antisocial action: Children's lies and behavior problems. *Journal of Moral Education*, 47, 432–450. https://doi.org/10.1080/03057240.2017.1405343

Lickenbrock, D. M., Braungart-Rieker, J. M., Ekas, N. V., Zentall, S. R., Oshio, T., & Planalp, E. M. (2013). Early temperament and attachment security with mothers and fathers as predictors of toddler compliance and noncompliance. *Infant and Child Development*, 22, 580–602. https://doi.org/10.1002/icd.1808

Loeber, R. (1991). Antisocial behavior: More enduring than changeable? *Journal of the American Academy of Child and Adolescent Psychiatry*, 30, 393–397.

Malti, T., & Dys, S. P. (2015). A developmental perspective on moral emotions. *Topoi*, 34, 453–459. https://doi.org/10.1007/s11245-013-9224-5

Meristo, M., Strid, K., & Surian, L. (2016). Preverbal infants' ability to encode the outcome of distributive actions. *Infancy*, 21, 353–372. https://doi.org/10.1111/infa.12124

Meristo, M., & Surian, L. (2014). Infants distinguish antisocial actions directed towards fair and unfair agents. *PLoS One*, 9, e110553. https://doi.org/10.1371/journal.pone.0110553

Miake-Lye, I. M., Hempel, S., Shanman, R., & Shekelle, P. G. (2016). What is an evidence map? A systematic review of published evidence maps and their definitions, methods, and products. *Systematic Reviews*, 5, 28. https://doi.org/10.1186/s13643-016-0204-x

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. *Development and Psychopathology*, 13, 355–375.

Morrell, J., & Murray, L. (2003). Parenting and the development of conduct disorder and hyperactive symptoms in childhood: A prospective longitudinal study from 2 months to 8 years. *Journal of Child Psychology and Psychiatry*, 44, 489–508. https://doi.org/10.1111/1469-7610.t01-1-00139

Mugno, A., Malloy, L. C., Waschbusch, D. A., Pelham, W. E., & Talwar, V. (2019). An experimental investigation of antisocial lie-telling among children with disruptive behavior disorders and typically-developing children. *Child Development*, 90, 774–789. https://doi.org/10.1111/cdev.12985

Narvaez, D., & Vaydich, J. L. (2008). Moral development and behavior under the spotlight of the neurobiological sciences. *Journal of Moral Education*, 37, 289–312. https://doi.org/10.1080/03057240802227478

Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition*, 108, 222–231. https://doi.org/10.1016/j.cognition.2007.12.003

Padilla-Walker, L. M., & Christensen, K. J. (2011). Empathy and self-regulation as mediators between parenting and adolescents' prosocial behavior toward strangers, friends, and family. *Journal of Research on Adolescence*, 21, 545–551. https://doi.org/10.1111/j.1532-7795.2010.00695.x

Patterson, G. R., DeGarmo, D. S., & Knutson, N. (2000). Hyperactive and antisocial behaviors: Comorbid or two points in the same process? *Development and Psychopathology*, 12, 91–106.

Paulus, M., Licata, M., Kristen, S., Thoermer, C., Woodward, A., & Sodian, B. (2015). Social understanding and self-regulation predict pre-schoolers' sharing with friends and disliked peers: A longitudinal study. *International Journal of Behavioral Development*, 39, 53–64. https://doi.org/10.1177/0165025414537923

Peters, M., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, 13, 141–146. https://doi.org/10.1097/XEB.0000000000000050

Poehlmann, J., Schwichtenberg, A. M., Hahn, E., Miller, K., Dilworth-Bart, J., Kaplan, D., & Maleck, S. (2012). Compliance, opposition, and behavior problems in toddlers born preterm or low birthweight. *Infant Mental Health Journal*, *33*, 34–44. https://doi.org/10.1002/imhj.20335

Shaw, D. S., Bell, R. Q., & Gilliom, M. (2000). A truly early starter model of antisocial behavior revisited. *Clinical Child and Family Psychology Review*, *3*, 155–172.

Smetana, J. G. (2013). Moral development: The social domain theory view. In P. D. Zelazo, (Ed.), *The Oxford handbook of developmental psychology: Body and mind*, (Vol. 1, pp. 832–863). Oxford: Oxford University Press.

Smetana, J. G., Metzger, A., Gettman, D. C., & Campione-Barr, N. (2006). Disclosure and secrecy in adolescent-parent relationships. *Child Development*, *77*, 201–217. https://doi.org/10.1111/j.1467-8624.2006.00865.x

Sodian, B., Licata, M., Kristen-Antonow, S., Paulus, M., Killen, M., & Woodward, A. (2016). Understanding of goals, beliefs, and desires predicts morally relevant theory of mind: A longitudinal investigation. *Child Development*, *87*, 1221–1232. https://doi.org/10.1111/cdev.12533

Sucharew, H., & Macaluso, M. (2019). Methods for research evidence synthesis: The scoping review approach. *Journal of Hospital Medicine*, *14*, 416–418. https://doi.org/10.12788/jhm.3248

Tafreshi, D., Thompson, J. J., & Racine, T. P. (2014). An analysis of the conceptual foundations of the infant preferential looking paradigm. *Human Development*, *57*, 222–240. https://doi.org/10.1159/000363487

Tan, E., Mikami, A. Y., & Hamlin, J. K. (2018). Do infant sociomoral evaluation and action studies predict preschool social and behavioral adjustment? *Journal of Experimental Child Psychology*, *176*, 39–54. https://doi.org/10.1016/j.jecp.2018.07.003

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, *58*, 345–372. https://doi.org/10.1146/annurev.psych.56.091103.070145

Thompson, R. A. (2012). Whither the preconventional child? Toward a life-span moral development theory. *Child Development Perspectives*, *6*, 423–429.

Turiel, E. (2015). Moral development. In W. F. Overton, P. C. M. Molenaar, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Theory and method* (pp. 484–522). Hoboken, NJ: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118963418.childpsy113

Van de Vondervoort, J. W., & Hamlin, J. K. (2016). Evidence for intuitive morality: Preverbal infants make sociomoral evaluations. *Child Development Perspectives*, *10*, 143–148. https://doi.org/10.1111/cdep.12175

Vaish, A., Carpenter, M., & Tomasello, M. (2009). Sympathy through affective perspective taking and its relation to prosocial behavior in toddlers. *Developmental Psychology*, *45*, 534–543. https://doi.org/10.1037/a0014322

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301–1303. https://doi.org/10.1126/science.1121448

Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*, 271–294. https://doi.org/10.1111/j.1532-7078.2007.tb00227.x

White, J. H., Moffit, T., Earls, F., & Robins, L. (1990). Preschool predictors of persistent conduct disorder and delinquency. *Criminology*, *28*, 443–454.

Wynn, K., & Bloom, P. (2014). The moral baby. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 435–453). New York: Psychology Press.

Wynn, K., Bloom, P., Jordan, A., Marshall, J., & Sheskin, M. (2018). Not noble savages after all: Limits to early altruism. *Current Directions in Psychological Science*, *27*, 3–8. https://doi.org/10.1177/0963721417734875

Ziv, T., & Sommerville, J. A. (2017). Developmental differences in infants' fairness expectations from 6 to 15 months of age. *Child Development*, *88*, 1930–1951. https://doi.org/10.1111/cdev.12674

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.