The missing link? Testing a schema account of unitization

Roni Tibon[1] · Andrea Greve[1] · Richard Henson[1]

Abstract

Unitization refers to the creation of a new unit from previously distinct items. The concept of unitization has been used to explain how novel pairings between items can be remembered without requiring recollection, by virtue of new, item-like representations that enable familiarity-based retrieval. We tested an alternative account of unitization – a schema account – which suggests that associations between items can be rapidly assimilated into a schema. We used a common operationalization of "unitization" as the difference between two unrelated words being linked by a definition, relative to two words being linked by a sentence, during an initial study phase. During the following relearning phase, a studied word was re-paired with a new word, either related or unrelated to the original associate from study. In a final test phase, memory for the relearned associations was tested. We hypothesized that, if unitized representations act like schemas, then we would observe some generalization to related words, such that memory would be better in the definition than sentence condition for related words, but not for unrelated words. Contrary to the schema hypothesis, evidence favored the null hypothesis of no difference between definition and sentence conditions for related words (Experiment 1), even when each cue was associated with multiple associates, indicating that the associations can be generalized (Experiment 2), or when the schematic information was explicitly re-activated during Relearning (Experiment 3). These results suggest that unitized associations do not generalize to accommodate new information, and therefore provide evidence against the schema account.

Keywords Unitization · Schema · Recollection · Familiarity · Recognition

Introduction

Did I put my car keys on the kitchen table? What's the name of this person? Did I ever see my friend wearing this outfit before? To answer these questions, which can be fundamental to our everyday lives, we need to be able to remember that two or more items were experienced conjointly. Unitization – the creation of a new unit from individual items – has been proposed as an effective strategy for remembering such episodic associations (e.g., Gobet et al., 2001; Graf & Schacter, 1989; Yonelinas, 1997; Yonelinas, Kroll, Dobbins, & Soltani, 1999). Although unitization effects on memory are well documented, the underlying cognitive mechanisms are only sparsely investigated in the current literature. The goal of the current study was to distinguish between two possible accounts of

unitization: a traditional "item account" and an alternative "schema account." We rely on a key difference between the two accounts – that the schema account, but not the item account, allows generalization from one association to another related one – and argue that if unitized associations do not display this critical property of schemas, then the item account of unitization remains most likely.

Memory for episodic associations can be assessed by presenting two items and asking whether they were previously presented together, a task termed "associative recognition." The widely supported dual-process theory of episodic memory posits that recognition comprises two separable processes: familiarity and recollection. Familiarity is a feeling of having encountered something or someone before, without retrieval of additional information, whereas recollection further provides contextual details about that encounter. Neurocognitive models of recognition memory claim that, while the hippocampus is necessary for recollection, familiarity is mainly associated with the perirhinal cortex (for review, see Yonelinas, 2002). Traditionally, it has been thought that, while recognition of single items can be supported by either recollection or familiarity, retrieval of newly learnt associations between arbitrary

✉ Roni Tibon
roni.tibon@mrc–cbu.cam.ac.uk

[1] MRC Cognition and Brain Sciences Unit, University of Cambridge, 5 Chaucer Road, Cambridge CB2 7EF, UK

items requires recollection (e.g., Donaldson & Rugg, 1998; Hockley & Consoli, 1999; Yonelinas, 1997). Subsequent findings, however, challenged this traditional view.

Quamme, Yonelinas, and Norman (2007) showed that amnestic patients with damage to the hippocampus and severe recollection deficits were nonetheless able to remember pairings of initially unrelated words following unitization. In their study, participants were given an arbitrary word pair such as CLOUD-LAWN in the context of either a definition (e.g., "a garden used for sky-gazing") or a sentence (e.g., "the _____ could be seen from the _____"). It was presumed that the former, but not the latter, creates a new unit that allows the two words to be encoded as a compound. The main finding was that although the patients struggled to remember nonunitized pairs, their memory for unitized pairs was relatively intact. This study indicates that hippocampal-based recollection might not be necessary for recognition of arbitrary pairings, and that unitization can enable familiarity-based retrieval on the basis of neocortical representations, for example in perirhinal cortex (see also D'Angelo, Kacollja, Rabin, Rosenbaum, & Ryan, 2015; Diana, Yonelinas, & Ranganath, 2010; Ryan, Moses, Barense, & Rosenbaum, 2013; for additional evidence from amnestic patients).

This notion was supported by a growing body of evidence (for review, see Mecklinger & Jäger, 2009; Yonelinas, Aly, Wang, & Koen, 2010), including intact memory for unitized associations in healthy older adults with decreased hippocampal functionality (Bastin, Diana, Simon, Collette, Yonelinas, & Salmon, 2013; Delhaye, Tibon, Gronau, Levy, & Bastin, 2017; Delhaye & Bastin, 2016; Memel & Ryan, 2017; Troyer, D'Souza, Vandermorris, & Murphy, 2011), behavioral studies showing that unitization affects familiarity but not recollection (e.g., Ahmad & Hockley, 2017; Diana, Yonelinas, & Ranganath, 2008; Parks & Yonelinas, 2015; Robey & Riggins, 2017; Shao, Opitz, Yang, and Weng, 2016; Tibon, Vakil, Goldstein, & Levy, 2012; Tu, Alty, & Diana, 2017), functional magnetic resonance imaging (fMRI) studies showing that retrieval of unitized information produces activation differences in regions associated with familiarity (Bader, Opitz, Reith, & Mecklinger, 2014; Diana et al., 2010; Ford, Verfaellie, & Giovanello, 2010; Haskins, Yonelinas, Quamme, & Ranganath, 2008; Memel & Ryan, 2017), and electrophysiological studies showing modulation of the well-documented FN400 ERP effect – the putative electrophysiological correlate of familiarity (reviewed by Mecklinger, 2000; Rugg & Curran, 2007; Wilding & Ranganath, 2011) – following unitization encoding (e.g., Bader, Mecklinger, Hoppstädter, & Meyer, 2010; Diana, Van den Boom, Yonelinas, & Ranganath, 2011; Guillaume & Etienne, 2015; Jäger, Mecklinger, & Kipp, 2006; Jäger, Mecklinger & Kliegel, 2010; Kamp, Bader, & Mecklinger, 2016; Rhodes & Donaldson, 2007; Rhodes & Donaldson, 2008; Tibon, Ben-Zvi, & Levy, 2014; Tibon, Gronau, Scheuplein,

Mecklinger, & Levy, 2014; Zheng, Li, Xiao, Broster, Jiang, & Xi, 2015).

The conventional account of these findings is that unitization results in the formation of a new representation, like that for a single item, rather than just strengthening associations between the original two item representations. In other words, according to this "item account," the two items become fully integrated or "chunked," so that they comprise a single unit, which does not necessarily bear any relationship to the constituent items (e.g., in the way that "cottage pie" has no direct semantic relationship with "cottage"). The core assumptions of the aforementioned studies are that: (1) familiarity *cannot* support retrieval of arbitrary episodic associations, and (2) item memory *can* be supported by familiarity. Therefore, if there is an indication of familiarity-based associative retrieval, then it is assumed that the to-be-associated items were fused together into a single-item representation. Nevertheless, this logic assumes that the "item" is the only cognitive construct that can be supported by familiarity. An alternative account would posit that unitization can generate another construct, which is not necessarily an item, but can still be retrieved via familiarity. One such alternative construct is a schema.

The notion of schemas in memory was introduced by Bartlett (1932), who observed that an organized structure of prior knowledge – a "schema" – can facilitate memory for new episodic information. Such schemas are stored in semantic memory and consist of chunks of knowledge about the world that have become activated in order to guide current processing (e.g., Anderson, 1981; Rumelhart, 1980). A recent framework (SLIMM: Schema-Linked Interactions between Medial prefrontal and Medial temporal lobes), which integrates schema theory with recent neuroscientific data, defines a schema as an activated network of interconnected neocortical representations (van Kesteren, Ruiter, Fernández, & Henson, 2012). Crucially, SLIMM proposes that a schema can support rapid acquisition of new knowledge that is congruent with that schema into the neocortex, without requiring the hippocampus. Given the evidence that unitized associations are encoded as hippocampus-independent neocortical representations (e.g., Quamme et al., 2007), a schema account of unitization should be considered.

A schema account would argue that rather than fusing distinct items into one unit, unitization rapidly converts arbitrary episodic links into meaningful semantic ones, thus establishing a schema. According to this, when two items (e.g., the words "CLOUD" and "LAWN") are presented together and accompanied by a semantic context that can apply to both (e.g., "a garden used for sky-gazing"), then even if the word-pair does not constitute a pre-existing concept, the semantic context provides a new structure that can accommodate both words. Critically, given that similar to item memory, semantic links are accessible via familiarity, i.e., a feeling of knowing (e.g., Kikyo, Ohki, & Miyashita, 2002; Tulving, 1972;

Yonelinas, 2002), the abovementioned studies are also consistent with a schema account. Furthermore, in addition to the core finding – that unitized associations can be retrieved via familiarity – which does not favor either account, other aspects of the previous results resonate with our proposed schema account. For example, the evidence that semantic memory remains preserved in amnestic patients (e.g., Vargha-Khadem, Gadian, Watkins, Connelly, Van Paesschen, & Mishkin, 1997), and in older age (e.g., Craik & Bosman, 1992), is consistent with the intact unitization effect found in these populations. Moreover, in fMRI studies, unitization was associated not only with perirhinal cortex, but also other neocortical regions that are believed to play a role in semantic/schema functions (e.g., medial prefrontal cortex, Bader et al., 2014; fusiform gyrus and superior temporal gyrus, Haskins et al., 2008). Indeed, when interpreting their findings of increased BOLD activation in the medial prefrontal cortex following unitization encoding, Bader et al. (2014) speculated that conditions that enable unitization render new information congruent with pre-existing knowledge, allowing rapid integration of that information. This post-hoc suggestion matches our proposed schema account. Finally, it is possible that what is observed in ERP studies of unitization is in fact a modulation of the N400 component (e.g., Kutas and Federmeier, 2011), which has been shown sensitive to schema congruent vs. incongruent information, and is difficult to distinguish from the FN400 (e.g., Paller, Voss, & Boehm, 2007).

An idea similar, but not identical, to our proposed schema account was considered previously by Parks and Yonelinas (2015), namely that unitization reflects semantic elaboration. These authors argued that to the extent that unitization creates a new unit, it should not improve recognition of the constituent items (i.e., improve associative recognition but not item recognition). In contrast, if unitization reflects semantic elaboration, it should benefit both associative and item recognition. The authors found that unitization did increase associative recognition but not item recognition (which remained unchanged), arguing against the semantic elaboration account. Others have argued that the item account predicts *decreased* familiarity of the constituent items, because only the unitized whole and not the constituent elements is readily available to familiarity at retrieval (Mayes, Montaldi, & Migo, 2007; Pilgrim, Murray, & Donaldson, 2012). Nevertheless, the schema account, unlike the semantic elaboration account, is also consistent with this suggestion of decreased familiarity. This is because schema theories typically argue that schemas represent commonalities, and therefore ignore item-specific details (e.g., Ghosh & Gilboa, 2014). Therefore, changes to item familiarity alone cannot be used to falsify our proposed schema account, and this should be tested by other means.

If current evidence is insufficient to distinguish item and schema accounts of unitization, how can they be tested? We suggest that a key property of the schema but not item account

is the potential for generalization. In its most basic form, a schema (S) is an abstract structure that allows two initially unrelated items (A and B) to become bound together, by virtue of links between A and S and between B and S. For example, in the presence of semantic links between "a garden for sky-gazing" (S) and CLOUD (A), and between "a garden for sky-gazing" (S) and LAWN (B), the novel association between CLOUD and LAWN (A-C) can be easily established. Then, if the schema (S) is reactivated later by a new pairing of MOON-LAWN (D-B), this should also allow rapid learning because MOON also relates to "a garden for sky-gazing," i.e. the schema enables generalization to new, related pairings (see Fig. 1). If the new pairing were however TEA-LAWN, then the lack of any semantic relationship between TEA and "a garden for sky-gazing" would not allow the schema to facilitate learning of TEA-LAWN. This schema account contrasts with the item account, which predicts that CLOUDLAWN becomes a new item representation that is not strongly linked to the original CLOUD and LAWN representations. Because of this unique (conjunctive) nature, no benefit is predicted for subsequent learning of the related MOON-LAWN pairing (i.e, learning should be no different from the unrelated TEA-LAWN pairing).

Note that this schema account is different from the semantic elaboration account refuted by Parks and Yonelinas (2015), in that linking a new pair of words (e.g, CLOUD and LAWN) into an abstract schema ("a garden used for sky-gazing") might entail elaboration of some aspects of the meanings of these words (e.g., to do with lying down and observing the sky), but is also likely to entail reduced elaboration of schema-irrelevant meanings (e.g., impending rain and cutting the grass), relative for example to the semantic elaboration entailed by the sentence condition ("the cloud could be seen from the lawn"). Note also that the schema account of generalization is more than simply semantic mediation, whereby representation of LAWN re-activates the previously paired associate CLOUD, and a pre-existing semantic link between CLOUD and MOON then allows CLOUD to "mediate" the new association MOON-LAWN. While this may occur, there is no obvious reason why activation of a semantic mediator should be more common in a definition condition than a sentence condition. Rather, we propose that an abstract structure, such as "a garden used for sky-gazing", is more likely to be generated in the definition condition than sentence condition, and only this abstract structure needs to be reactivated in order to facilitate new learning of MOON-LAWN (i.e, act as a schema), without necessarily re-activating the original study pair (CLOUD-LAWN).

We therefore performed three experiments along the lines above, in order to distinguish the item and schema accounts of unitization.
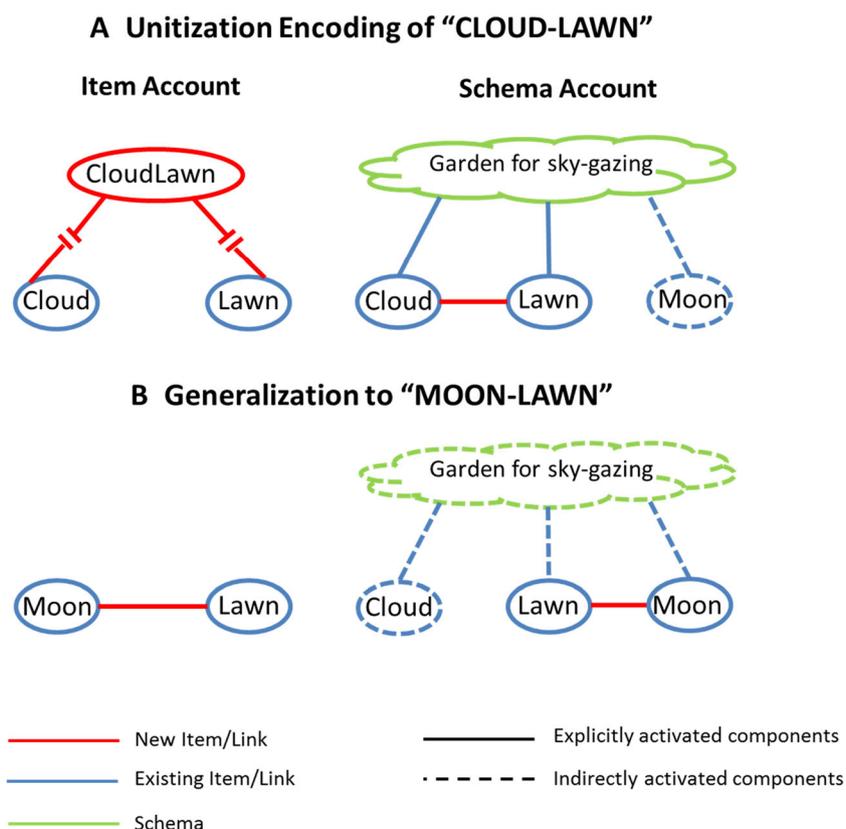
A Unitization Encoding of "CLOUD-LAWN"

Item Account

Schema Account

CloudLawn

Garden for sky-gazing

Cloud   Lawn

Cloud — Lawn   Moon

B Generalization to "MOON-LAWN"

Garden for sky-gazing

Moon — Lawn   Cloud   Lawn — Moon

— New Item/Link   — Explicitly activated components

— Existing Item/Link   - - - Indirectly activated components

— Schema

Fig. 1 Schematic depiction of unitization encoding (**top**) and generalization (**bottom**) under an item account (**left**) and a schema account (**right**). Blue coloring indicates items and links that already exist in the network (we use the term "existing links" to refer to items that similarly fit with the definition). Red coloring indicates new links and items that are learned during encoding or generalization. Solid lines indicate components that are explicitly activated during encoding or generalization, while dashed lines indicate components in the network that are not explicitly presented but can be activated indirectly (e.g., via pre-existing links). Interrupted red lines (item account) represent temporary links, which only remain active until the "fused" representation is created. Green cloud (schema account) represents a schema that can bind items together

Experiment 1

The basic design of Experiment 1 is illustrated in Fig. 2. Following Bader et al. (2010) and Quamme et al. (2007), we operationalized "unitization" as the difference between two words being linked by a definition (e.g., CLOUD/LAWN, *A garden used for sky-gazing*), relative to two words being linked by a sentence (e.g., CLOUD/LAWN, *The _____ could be seen from the _____*). This manipulation of encoding instructions was done during an initial study phase. We crossed these two conditions with the nature of a second word-pair presented in a subsequent relearning phase, which was either the same pairing as in initial learning (repeat condition, e.g., CLOUD-LAWN), a related pairing (e.g., MOON-LAWN) or unrelated pairing (e.g., TEA-LAWN). Then in the final test phase, participants selected the word from the relearning phase (i.e., MOON) that had been paired with a cue word (e.g., LAWN), from one of four alternatives (see Methods for precise description of foils). We hypothesized that, if schemas underline unitization, then there should be some generalization to related ideas, such that there would be a greater advantage of definition over sentence conditions for related than for unrelated conditions. We further expected this effect to be accompanied by increased erroneous selection of new related rather than unrelated lures (e.g., STAR) in the definition over sentence conditions. The purpose of the repeat condition was to check that we replicate the basic advantage of definition versus sentence conditions (i.e., the unitization advantage) that has been reported previously.

Methods

Participants Twenty-four Cambridge community members participated in the experiment (21 females, mean age 27.3 years, $SD = 5.77$). In Parks and Yonelinas (2015; Experiment 1), a similar stimulus dataset and a similar (encoding) task were used in a between-participant design, with 22 participants in each encoding condition. Based on the data presented in Fig. 1 of their paper, we estimated $M = 0.71$, $SD = 0.328$ in the definition condition, and $M = 0.16$, $SD = 0.169$ in the sentence condition. Cohen's d was therefore estimated to be 2.1, which meant that testing six participants
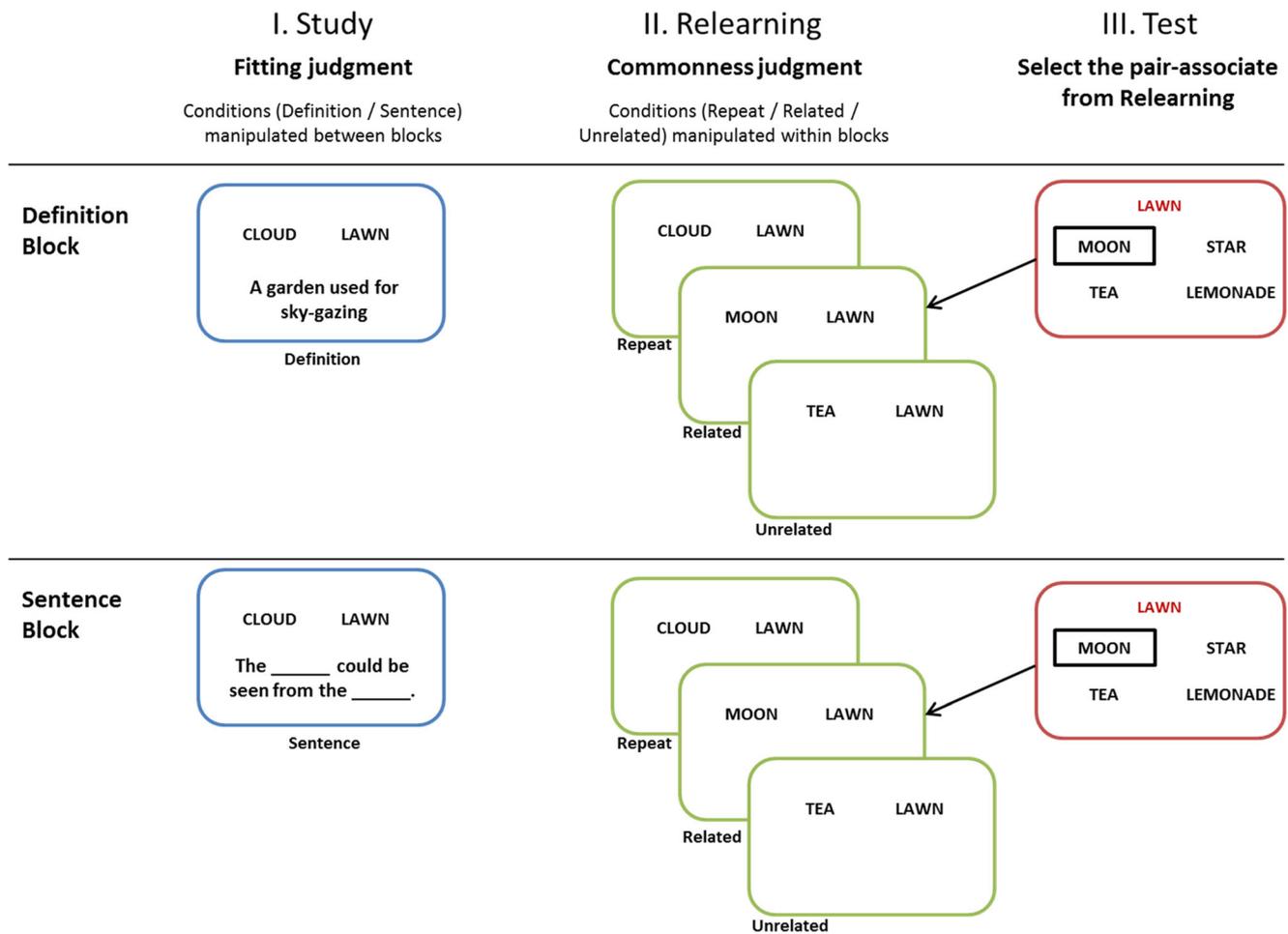
I. Study
Fitting judgment
Conditions (Definition / Sentence) manipulated between blocks

II. Relearning
Commonness judgment
Conditions (Repeat / Related / Unrelated) manipulated within blocks

III. Test
Select the pair-associate from Relearning

Fig. 2 Illustration of the experimental design of Experiment 1. Definition and sentence blocks only differed at study phase. The test trial in the example corresponds to the related trial in the relearning phase. The target response (MOON) is highlighted. Other response types are old-different (TEA), new-related (STAR), and new-unrelated (LEMONADE). Repetition of word pairs in the figure is for illustrative purposes: during the task, each word-pair was only presented once under definition / sentence encoding and once under repeat / related / unrelated relearning

should be sufficient in order to obtain statistical power of 95%. Nevertheless, we had to test at least 24 participants to achieve a full counterbalance of our experimental design, which according to the above calculations would provide a power > 99%.

Two participants were excluded from the study due to very fast averaged response times in some conditions (< 750 ms, while minimum averaged RT after excluding these participants was 1,923 ms). These participants were replaced with two others, to keep the dataset counterbalanced. For one participant, a software crash caused 17 relearning trials to be repeated, but these trials were later excluded. For all experiments, participants were recruited from the volunteer panel of the MRC Cognition and Brain Science Unit, were native English speakers, had normal or corrected-to-normal vision, and were never diagnosed with attention-deficit hyperactivity disorder (ADHD), dyslexia, or any other developmental or learning disabilities. They all provided informed consent and

received monetary compensation for participation, as approved by a local ethics committee (Cambridge Psychological Research Ethics Committee reference 2005.08).

Materials Initial study materials included 288 entries. Each entry was comprised of semantically unrelated word pairs and their corresponding definitions and sentences that were used in Haskins et al. (2008). For example, one of the entries in this database included the word-pair CLOUD LAWN, which corresponded to the definition "A garden used for sky-gazing" and to the sentence "The _____ could be seen from the _____". For each entry, we designated one word of each word-pair to serve as the cue (thus remaining constant across phases), and the other to serve as its associate. For 50% of the entries in the database the first word was assigned to be the cue-word, and for the remaining 50% the second word was the cue word. This assignment remained constant across

experimental lists. We then added to each entry two additional associates (also unrelated to the cue) that would fit the definitions and the sentences to a similar degree as the original associate. For example, for the example described above, the cue word LAWN, which was paired with CLOUD (its original associate), was also paired with MOON and STAR that fit the definition and sentence in the same way as the word CLOUD. In some cases, minor modifications were made to the original entries (from Haskins et al., 2008) to allow more words to fit with the definitions/sentences (e.g., for the word pair COFFEE PERSON, the definition "A vendor of caffeinated beverages" was modified to "A vendor of beverages," in order to also fit TEA and LEMONADE as alternatives to COFFEE). The full database was divided into 24 sub-lists, with 12 entries in each sub-list. In each sub-list, each entry was pseudo-randomly associated with the following entry (the last entry was associated with the first one), in order to add to this entry two other words that are unrelated, i.e., do not fit with the definitions/sentences (e.g., CLOUD was further linked with TEA and LEMONADE). Thus, each associate (CLOUD) was linked with two related alternatives (MOON and STAR) and two unrelated alternatives (TEA and LEMONADE).

The fit of the original associate and its alternatives with the definitions and sentences were assessed by eight raters (four for the definitions and four for the sentences) who did not take part in the main experiments and were all native English speakers. They were asked to decide if the related words similarly fit the definitions/sentences or not, and if the unrelated words similarly misfit the definitions/sentences. If not, they were asked to come up with a suggestion that they think would be more appropriate. Words marked as unfit (for related words) or as fit (for unrelated words) by more than one rater were replaced with the words suggested. These new matches were verified by two other independent judges.

Design and procedure The experiment was run using E-Prime 2 (Psychology Software Tools). Participants sat in front of a computer monitor within a comfortable viewing distance. All verbal stimuli (except for cue-word at test, see below), were displayed in Courier New 18-point font, in black against a gray background. Our design included three experimental phases: study, relearning and test, with an additional distracting task (counting backwards for 30 s) succeeding study and relearning phases.

The study phase included two conditions: during the *definition encoding* condition, participants were instructed to rate the pair as a whole on a scale ranging from 1 ("very badly") to 4 ("very well") according to how well the definition combined the meanings of the two words into a sensible compound. In the *sentence encoding* condition, participants were instructed to rate how well the two words fitted into the sentence frame. Study trials began with a 500-ms fixation cross. Next, a word-pair (e.g., CLOUD LAWN) and a definition/sentence were

displayed for 3,500 ms one above the other, slightly above and below central vision. Sentence frames were constructed with two blank spaces, where the first item was intended to fit into the first space, and the second item into the second space. This was followed by a 100-ms blank screen, which was then replaced by a screen displaying a question mark. Participants were instructed to respond when the question mark appears, via the keys '1', '2', '3' and '4' on the left side of the keyboard. Response time was self-paced, and triggered the onset of the next trial.

During the relearning phase, participants were presented with a word-pair, and were asked to judge which of the two words is more frequent in their daily life, using the left arrow to indicate the left word, or the right arrow to indicate the right word. This phase included three relearning conditions: in the *repeat* condition, word-pairs were identical to pairs presented at study (CLOUD LAWN). In the *related* condition, cue-words remained the same, but associates were replaced with a related alternative (MOON LAWN). In the *unrelated* condition, cue-words remained the same, but associates were replaced with an unrelated alternative (TEA LAWN). The relearning task (frequency judgment) remained the same across these three conditions. Relearning trials began with a 500-ms fixation cross. Next, a word-pair was presented for 3,500 ms, slightly above central vision. As in the study phase, this was followed by a 100-ms blank screen, which was replaced by a response screen displaying a question mark where participants were asked to provide their responses.

Finally, at the test phase, participants were presented with a cue word (LAWN) shown in red at the top of the screen. Below it, four words appeared. One was its paired-associate from the relearning phase (target, correct response: CLOUD in the repeated condition, MOON in the related condition and TEA in the unrelated condition). The second was a word that was presented at relearning, but was paired with a different cue (old-different, e.g., TEA when target is MOON and TEA was presented in a different relearning trial). The other two lures were new words that were not presented before. One was the related alternative (new-related, e.g., STAR when target is MOON), and the remaining one was the unrelated alternative (new-unrelated, e.g., LEMONADE when target is MOON and TEA was presented in a different relearning trial). Participants were asked to indicate which word was associated with the cue-word *at relearning*. Test trials began with a 500-ms fixation cross, followed by the presentation of the cue-word and four alternative forced choice (AFC). Participants were asked to press "R" to select the top-left word, "C" to select bottom-left, "U" to select top-right, and "M" to select bottom-right. Responses were self-paced and triggered the appearance of a 100-ms blank screen, followed by the next trial.

Encoding conditions were presented in different blocks, with block order (either definition - sentence – sentence – definition or sentence – definition – definition – sentence)

counterbalanced across participants. This factor did not interact with any of the experimental conditions, and therefore was not investigated further. Relearning conditions were randomized within and between blocks. Each block included 24 word-pairs in each relearning condition (total of 72 per block). Overall, the experiment included four blocks, each containing all three experimental phases.

Results

All statistical analyses were performed using R version 3.4.1 and RStudio Version 1.0153, with packages BayesFactor (Morey & Rouder, 2015), ez (Lawrence, 2016), and psych (Revelle, 2017). In addition to classical hypothesis testing, we also employed Bayesian statistics throughout the entire study, to allow interpretation of null results. Unlike classical hypothesis testing, in which the p-value is calculated based on the assumption of drawing a hypothetical infinite number of samples (i.e., sampling distribution), the Bayesian approach makes direct probability statements about the parameters using the observed sample. Under this approach, a Bayes factor (BF) is used to quantify the support for one model (e.g., the alternate hypothesis) over another (e.g., the null model). It is generally agreed that a model is favored when a Bayes factor greater than 3 is obtained (Rouder, Speckman, Sun, Morey, & Geoffrey, 2009).

Performance during study To test for differences in relatedness rating at study, we used a two-sided pairwise t-test and a Bayesian t-test with a Cauchy prior scaled at sqrt(2)/2 (medium scaling), contrasting averaged rating scores in the definition versus sentence condition. This revealed a significant difference, $t(23) = -2.93$, $p = .008$, 95%, CI [-.51, -.09]; BF(h1) = 6.17, credential interval (CrI) [-.45, -.1], with word-pairs in the sentence condition ($M = 2.95$, $SD = 0.5$) being judged as more related than in the definition condition ($M = 5.65$, $SD = 0.49$).

We further conducted a logistic regression analysis, to ensure that these differences in relatedness ratings during study do not interact with accuracy at test. This analysis was set to predict retrieval outcome in each trial (0 – unsuccessful; 1 – successful) using relatedness rating at study (1 – 4). Coefficient values for the predictor were extracted separately for each participant in each encoding condition. A one-sample t-test, contrasting the mean value of these coefficients against zero, separately for each encoding condition, did not reveal any significant differences for definition encoding, $t(23) = .58$, $p = .57$, CI [-.06, .1], or for sentence encoding: $t(23) = -.02$, $p = .98$, CI [-.09, .09]. Furthermore, Bayesian t-tests showed evidence in favor of the null hypothesis of no difference between coefficients for definition encoding, BF(h0) = 4, CrI [-.04, -.07], and for sentence encoding, BF(h0) = 4.66, CrI [-.07, .07]. These results suggest that the ability to accurately retrieve pair-associates during the retrieval phase is not predicted by ratings at study.

Test performance: 4AFC Accuracy rates at test (defined as % correct responses; chance performance is 25%) in the various encoding and relearning conditions are shown in Fig. 3. To check that we replicate the basic unitization effect, we contrasted accuracy rates at retrieval in the repeat condition for definition versus sentence encoding using a one-sided paired-sample t-test and a Bayesian t-test. These revealed that memory following definition encoding was significantly better than following sentence encoding, $t(23) = 2.91$, $p = .004$, CI [.02, .09]; BF(h1) = 11.69, CrI [.02, .08], consistent with the unitization effect.

Next, we tested our hypothesized schema account of unitization, which we expressed as two predictions. First, our main prediction was a greater advantage of definition over sentence encoding conditions for related than for unrelated relearning conditions. To test this prediction, we used a 2 X 2 repeated measures ANOVA with encoding condition (definition, sentence) and relearning condition (related, unrelated) as repeated factors. This analysis revealed a significant effect of encoding condition, $F(1, 23) = 5.37$, $p = .029$, $^2G = .008$ (though stemming mostly from differences in the unrelated condition; see Fig. 3), with better memory following definition encoding compared to sentence encoding. The effect of relearning condition, $F(1, 23) = 1.18$, $p = .29$, $^2G = .003$, and the interaction between the two factors, $F(1, 23) = .46$, $p = .5$, $^2G < .001$, were not significant. In other words, our main prediction was not confirmed, and in fact, a numerical trend suggested the opposite pattern of a greater difference between related than unrelated in the sentence condition.

In addition, we used a Bayesian analysis to allow interpretation of null effects. For each participant, we calculated the difference in accuracy rates for definition versus sentence encoding, separately for related and unrelated relearning. We then used a one-sided Bayesian t-test to compare two hypotheses: that the difference between encoding conditions does not differ for related and unrelated relearning conditions (the so-called "null hypothesis"), or that this difference is bigger for related relearning condition versus unrelated relearning condition. This analysis supported the null model, with a Bayes factor of BF(h0) = 7.23, CrI [-.07, .03]. The data thus provide moderate evidence against the hypothesis that memory is better for related versus unrelated information following definition but not sentence encoding.

Our second prediction was increased selection of new-related lures in the definition over sentence encoding conditions for related than for unrelated relearning conditions. The error data are summarized in Table 1; the proportions of
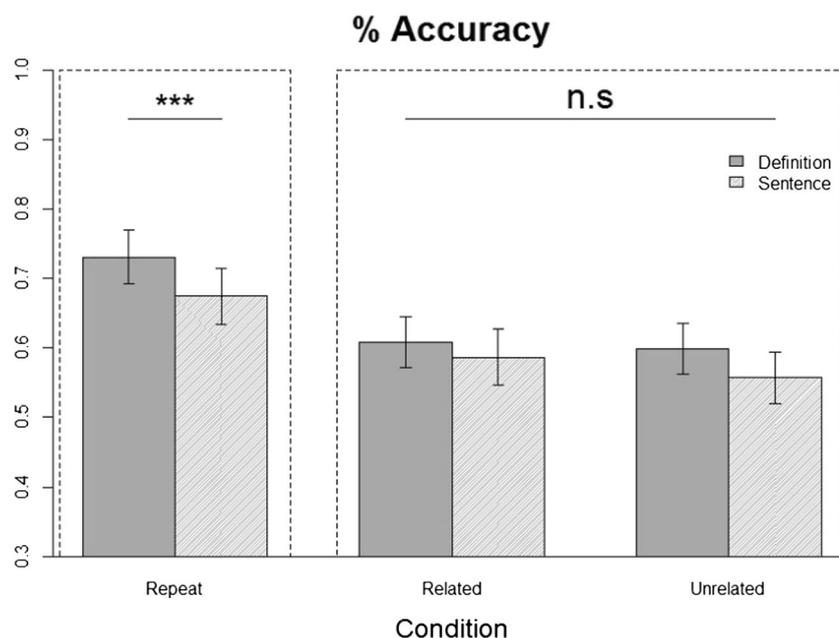
Fig. 3 Experiment 1: Proportion of correct responses at test, in the various encoding and relearning conditions. The two leftmost bars represent accuracy in the repeat (control) condition. Performance for repeated trials was significantly better for the definition than sentence conditions, consistent with the basic unitisation effect (***p < .005). The four bars on the right represent accuracy in the related and unrelated conditions. The predicted interaction between study condition and relearning condition was not significant (n.s). Error bars represent SEs for each condition separately

new-related distractors are shown in Table 2. These proportions were calculated by dividing the number of new-related responses by the total number of new responses (new-related / [new-related + new-unrelated]). Three participants who did not provide any new responses (thereby resulting in a zero denominator) were removed from this analysis. These proportions were subjected to a 2 X 2 ANOVA analysis with encoding condition and relearning condition (related, unrelated) as repeated factors, which revealed a significant effect of relearning condition, with increased selection of new-related distractors in the related condition, $F(1, 20) = 11.48$, $p = .003$, $^{2}G = .13$. The effect of encoding condition, $F(1, 20) = .08$, $p = .77$, $^{2}G < .001$, and the encoding-by-relearning interaction, $F(1, 20) = .29$, $p = .6$, $^{2}G = .002$, were not significant. A Bayesian t-test for the difference between definition and sentence encoding in proportion of new-related distractors for related versus unrelated relearning, revealed anecdotal evidence in support of the null hypothesis, BF(h0) = 2.79, CrI [-.08 .15].

Discussion

The results of Experiment 1 do not support a schema account, because unitized associations did not generalize to learning of new related information. We did not find any greater advantage of definition versus sentence encoding (our operationalization of unitization) when learning new related information than when learning new unrelated information, neither in accuracy rates nor in the tendency to erroneously select related lures. This is despite the fact that we did see an advantage of definition versus sentence encoding in the repeat condition, supporting the previously reported unitization effect.

It is possible, however, that unitized associations do generalize to accommodate new information, but only under certain conditions. One factor that might promote such generalization is the extent to which unitized information is perceived as generalizable. That is, if the same cue were associated at encoding with multiple related associates, indicating a more

Table 1 Error data for Experiment 1. Data represent % of errors made to each lure type. Standard deviations are given in parentheses

	Repeat			Related			Unrelated		
	Old-Different	New-Related	New-Unrelated	Old-Different	New-Related	New-Unrelated	Old-Different	New-Related	New-Unrelated
Definition	52.4 (19.7)	29.4 (21.2)	18.2 (13.9)	60.7 (15.3)	24.9 (10.6)	14.4 (9.7)	66.0 (13.7)	15.6 (11.0)	18.4 (7.9)
Sentence	61.4 (20.6)	21.7 (14.5)	16.9 (12.8)	60.5 (17.6)	21.3 (9.4)	18.2 (11.9)	59.6 (16.4)	20.0 (9.1)	20.4 (11.0)

Table 2 Proportion of new-related distractors in Experiment 1, for n=21 participants, calculated by dividing the number of new-related responses by overall number of new responses

	Repeat	Related	Unrelated
Definition	57.8 (24.9)	61.4 (21.8)	44.3 (23.3)
Sentence	55.6 (23.9)	60.5 (18.2)	47.4 (15.5)

general relationship (schema), would this encourage further generalization during relearning? Experiment 2 addressed this possibility.

Experiment 2

In Experiment 2, we added a second study round, in which each learned cue (e.g., the cue LAWN, which was initially studied with the target CLOUD) was presented again with the same definition/sentence, but with a different associate (e.g., PLANET) that similarly fit the definition/sentence. We hypothesized that, if the second study round indicates generalizability of the unitized association, then we would expect further generalization to related ideas. As in Experiment 1, we tested this prediction of the schema account in terms of the interaction between encoding condition and relearning condition, i.e., the prediction of a greater advantage of definition over sentence conditions for related than for unrelated conditions, and increased erroneous selection of new-related lures in the definition over sentence conditions for related than for unrelated conditions. Because of the additional study phase in each block, and in order to keep total experimental time similar to Experiment 1, Experiment 2 dropped the repeat (control) condition of Experiment 1.

Methods

Participants Sixteen Cambridge community members participated in the experiment (11 females, mean age 27.56 years, $SD = 7.11$).

Materials Study materials included the same stimuli database used for Experiment 1. However, in order to add an additional study trial to indicate generalization, we added another associate to each of the 288 entries, which would fit the definitions and the sentences to a similar degree as the other related associates. For example, the entry containing CLOUD LAWN was also paired with PLANET that fit the definition and sentence in a similar way to the word MOON LAWN. The fit of the new alternatives with the definitions and sentences was assessed by two independent judges who did not take part in the main experiments. Words marked as unfit were replaced and verified by the same two judges.

The new experimental design, which required two study rounds, meant that each block would take longer to complete compared to Experiment 1. Therefore, in order to keep total length roughly the same as Experiment 1, Experiment 2 did not include the repeat conditions (see design below) and stimuli database included 196 entries, randomly selected from the full 288-entries database.

Design and procedure The experimental design was similar to Experiment 1. Study phase included the same two encoding conditions (*definition* and *sentence* encoding), and relearning phase included the *related* and the *unrelated* relearning conditions.

The procedure followed that of Experiment 1, except for one modification: in each block the initial study phase was followed by a generalization phase, in which study trials were repeated but with cue-words now paired with the new alternatives (see Materials section above). Participants were not informed when the generalization phase started. They were informed, however, that the pairs will be repeated with some modifications, but were asked to focus on rating the fit for the pair as it currently appears. Procedures for relearning and test phase were identical to Experiment 1.

Encoding conditions were presented in different blocks, with block order counter-balanced across participants. Unlike in Experiment 1, however, in Experiment 2 this factor did interact with some of the experimental conditions, and was therefore included as a covariate factor in the ANOVAs for the test phase. Relearning conditions were randomized within blocks. Each block included 24 word-pairs in each relearning condition (total of 48 per block). Overall, the experiment included four blocks, each containing all four experimental phases (study, generalization, relearning, and test).

Results

Performance during study A pairwise t-test and a Bayesian t-test, contrasting averaged rating scores in the definition versus sentence condition, revealed a significant difference, $t(23) = -6.65$, $p < .001$, CI [-.79, -.41]; BF(h1) = 2782.39, CrI [-.73, -.4] with word-pairs in the sentence condition ($M = 3.13$, $SD = 0.44$) being judged as more related than in the definition condition ($M = 2.53$, $SD = 0.53$). We used the same procedure described for Experiment 1 above, to ensure that these differences in relatedness ratings during study did not interact with successful retrieval. Once again, one-sample t-tests and Bayesian t-tests contrasting Beta values against zero did not reveal significant differences, neither for definition encoding, $t(15) = 1.46$, $p = .17$, CI [-.05, .26], nor for sentence encoding, $t(15) = -1.02$, $p = .32$, CI [-.34, .12], and the Bayesian t-test showed anecdotal evidence in favor of the null hypothesis for both definitions, BF(h0) = 1.61, CrI [-.02, .2], and sentence encoding, BF(h0) = 2.49, CrI [-.26, .07].

Test performance: 4AFC Accuracy rates at test are shown in Fig. 4. To test the predictions of the schema account, we used the same analyses performed for Experiment 1, but also included block order as a between-subject covariate factor. For the first prediction, the 2 X 2 repeated measures ANCOVA with encoding and relearning conditions as repeated factors, and with block order as covariate, did not reveal any significant results: $F(1, 15) = .6$, $p = .45$, $^2G = .004$ for encoding condition; $F(1, 15) = 3.79$, $p = .07$, $^2G = .017$ for relearning condition; $F(1, 15) = .54$, $p = .47$, $^2G = .002$ for the interaction. The lack of interaction between encoding and relearning conditions was further supported by our Bayesian comparison of the null against the alternate hypothesis, revealing preference for the null hypothesis of no advantage for definition over sentences for related versus unrelated relearning, BF(h0) = 6.18, CrI [-.08, .04]. Thus, there was no evidence of greater advantage of definition over sentence encoding for related than for unrelated relearning condition

The error data are summarized in Tables 3 and 4. For our second prediction, the 2 X 2 ANCOVA for the proportion of new-related distractors, with encoding and relearning conditions as repeated factors, and with block order as covariate, revealed a significant effect of relearning condition, with increased selection of new-related distractors in the related condition, $F(1, 15) = 22.2$, $p < .001$, $^2G = .14$ but no effect of encoding condition, $F(1, 15) = .1$, $p = .76$, $^2G = .003$. Additionally, a significant interaction between encoding and relearning condition emerged, $F(1, 15) = 6.07$, $p = .026$, $^2G = .05$. Nevertheless, one-sided t-tests, performed separately for definition and sentence encoding, revealed that this interaction stemmed from increased selection of new-related distractors in the related versus unrelated relearning condition for items in the sentence encoding condition, $t(15) = 4.07$, $p = .001$, CI [.15, .39], but not in the definition encoding condition, $t(15) = 1.89$, $p = .08$, CI [-.01, .14], and therefore was in the opposite direction to our prediction. This was supported by our Bayesian analysis, which showed strong evidence in favor of the null hypothesis, BF(h0) = 11.12, CrI [-.03, .04], over the alternate hypothesis of greater difference between encoding conditions in the related than the unrelated relearning condition.

To summarize, definition encoding was not associated with better learning of related words, nor with increased selection of related new lures at test. Our hypothesis that, following generalized encoding, unitized associations would form a schema into which new information can be easily integrated, was not confirmed by Experiment 2 either.

Discussion

As with Experiment 1, the results of Experiments 2 supported the null hypothesis, suggesting that unitized associations do not generalize to facilitate learning of new, related

information. This is even though Experiment 2 included an additional study phase, which was intended to indicate to participants that the learned information was generalizable. However, in Experiments 1 and 2, the relearning phase involved a frequency judgment task, which can be accomplished without referring to the information that was learned during study. It is possible that in order for unitized associations to generalize, they need to be activated when new information is learned. In our final attempt to promote generalization, we designed a task in which the associated information from the study phase was explicitly re-activated during relearning.

Experiment 3

In Experiment 3, participants completed a memory task during the relearning phase, in which they were asked to identify old and new pairings (i.e., whether or not from the study phase), thereby activating memories from the study phase during the relearning phase. If re-activation promotes generalization of unitized associations, then we have the same predictions as in Experiments 1 and 2, i.e., greater accuracy rates at test, as well as increased erroneous selection of new-related lures, for definition over sentence encoding when related versus unrelated information is presented at relearning.

The inclusion of a memory task at relearning also allowed us to inquire whether there was an indication for generalization in this phase too. We predicted that if generalization occurs during relearning, then the definition encoding would increase the number of erroneous responses relative to sentence encoding, in which participants were more likely to classify newly learned associations as "old" for related than for unrelated pairs.

Finally, as in Experiment 1, we included a repeat condition as a control, to check that definition versus sentence encoding produced the same advantage associated with unitization. Moreover, we expected this advantage to emerge during relearning as well as the test phase.

Methods

Participants Twenty-four Cambridge community members participated in the experiment (20 females, mean age 23.96 years, $SD = 5.77$).

Design and procedure Materials and design were the same as in Experiment 1: the study phase included two conditions (*definition* and *sentence* encoding); the re-learning phase included three conditions (*repeat*, *related*, and *unrelated*); and the test phase included a 4AFC retrieval task. Block order in this experiment did not interact with experimental conditions, and was not investigated further.
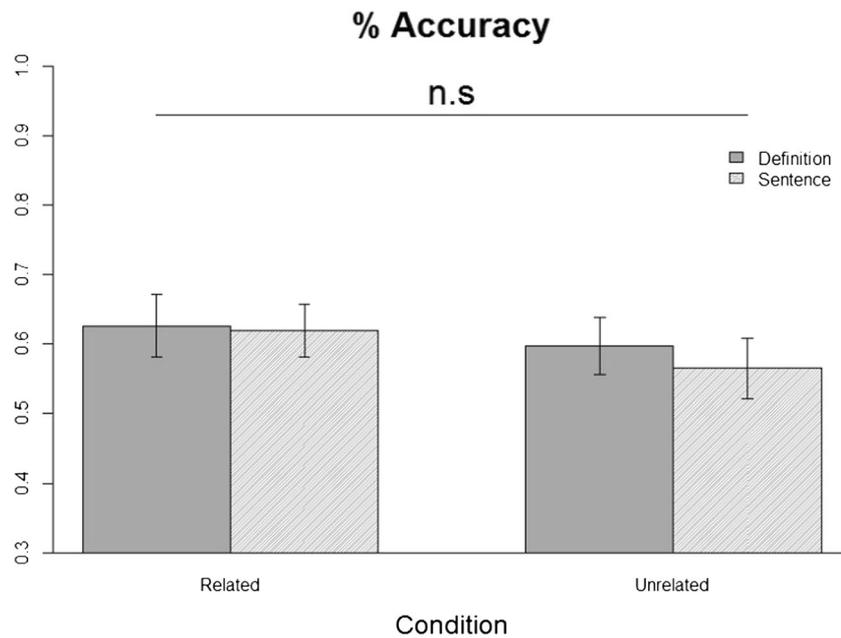
% Accuracy

Fig. 4 Experiment 2: Proportion of correct responses at test, in the various encoding and relearning conditions. Bars represent accuracy in the related and unrelated conditions. The predicted interaction between study condition and relearning condition was not significant (n.s.). Error bars represent SEs for each condition separately

During the relearning phase, participants had to complete a task different from the other two experiments. They were told that they will see word pairs, and were asked to indicate which they saw before in the study phase. They were asked to press a key with their index finger if they only saw the left word in the previous phase (an old-new pair), with their ring finger if they only saw the right word (a new-old pair), and with their middle finger if they saw both words together (old pairing). Because for half the pairs the first word was assigned to be the cue-word (which remained fixed), and for the other half the second word was the cue word (see Materials section for Experiment 1), repeated pairs were to receive old-pairing responses, while related and unrelated pairs were to receive either old-new or new-old responses. Participants were further instructed that to aid their memory, they should try to bring to mind the definitions or sentences from the previous phase.

Results

Performance during study As with the other experiments, a pairwise t-test and a Bayesian t-test revealed a significant difference, $t(24) = -6.45$, $p < .001$, CI [-.65, -.33]; BF(h1) = 13246.91, CrI [-.61, -.34], with word-pairs in the sentence condition ($M = 3.18$, $SD = .32$) being judged as more related than in the definition condition ($M = 2.68$, $SD = .46$). Once again, one-sample t-tests and Bayesian t-tests contrasting Beta values against zero did not reveal significant differences between successful and unsuccessful retrieval trials in their relatedness ratings at study, for neither definition encoding, $t(23) = .9$, $p = .37$, CI [-.06, .17]; BF(h0) = 3.22, CrI [-.04, .13], nor for sentence encoding $t(23) = 1.79$, $p = .09$, CI [-.02, .25]; BF(h0) = 1.18, CrI [.002, .2].

Performance during relearning We replicated once again the basic unitization effect, this time during the relearning phase: we contrasted accuracy rates at relearning following definition versus sentence encoding in the repeat condition using a one-sided paired-sample t-test and a Bayesian t-test. This revealed that definition ($M = 82$, $SD = .11$) encoding resulted in significantly better memory than sentence encoding ($M = 75$, $SD = .12$), $t(23) = 3.73$, $p < .001$, CI [.04, .11]; BF(h1) = 64.6, CrI [.03, .1].

Table 3 Error data for Experiment 2. Data represent % of errors that made to each lure type. Standard deviations are given in parentheses

Condition	Related			Unrelated		
	Old-Different	New-Related	New-Unrelated	Old-Different	New-Related	New-Unrelated
Definition	55.5 (15.4)	30.4 (15.4)	14.1 (9.1)	61.7 (16.3)	21.7 (10.9)	16.7 (9.1)
Sentence	66.0 (15.3)	23.6 (10.6)	10.4 (7.8)	62.9 (16.4)	17.7 (9.3)	19.4 (10.2)

Table 4 Proportion of new-related distractors in Experiment 2

	Related	Unrelated
Definition	66.0 (25.1)	58.8 (22.8)
Sentence	73.6 (18.3)	46.6 (18.0)

We further tested our prediction that a schema account would be indicated by greater number of erroneous "old" responses to new pairs in definition over sentence encoding for related than for unrelated relearning. A 2 X 2 repeated measures ANOVA and the corresponding Bayesian analysis, revealed a significant effect of relearning condition, $F(1, 23) = 55.75, p < .001, {}^2G = .22$, but no significant effect of encoding condition, $F(1, 23) = 1.42, p = .25, {}^2G = .008$, and no interaction between the two factors, $F(1, 23) = 2.04, p = .17, {}^2G = .006$; BF(h0) = 10.23, CrI [-.04, .004].

Performance at test: 4AFC Accuracy rates are shown in Fig. 5. Like in Experiment 1, we replicated the basic unitization effect at test, with increased accuracy rates following definition versus sentence encoding in the repeat condition, $t(23) = 2.91, p = .004$, CI [.02, 0.8]; BF(h1) = 11.7, CrI [.02, .07]. A 2 X 2 repeated measures ANOVA, set to test for greater advantage of definition over sentence encoding for related than for unrelated relearning, revealed a significant main effect of relearning condition, $F(1, 23) = 180.26, p < .001, {}^2G = .17$, but no main effect for encoding condition, $F(1, 23) = 3.43, p = .08, {}^2G = .01$, and no interaction, $F(1, 23) = .03, p = .87, {}^2G < .001$. This lack of interaction was further supported by our Bayesian analysis, confirming support for the null hypothesis of no difference, BF(h0) = 5.24, CI [-.06, .05].

The error data are summarized in Tables 5 and 6. To test our prediction of increased selection of new-related lures in the definition over sentence encoding for related than for unrelated relearning, we excluded seven participants who did not provide any responses of these response types. The 2 X 2 ANOVA for the proportion of new-related distractors revealed a significant effect of relearning condition, with increased selection of new-related distractors in the related condition, $F(1, 16) = 9.82, p = .006, {}^2G = .08$, but no effect of encoding condition, $F(1, 16) = 4.04, p = .06, {}^2G = .06$, and no interaction between these factors, $F(1, 16) = 1.13, p = .3, {}^2G = .02$. Despite the numerical trend in the predicted direction, the lack of interaction was somewhat confirmed by the Bayesian analysis, which showed anecdotal evidence in favor of the null hypothesis, BF(h0) = 1.48, CrI [-.06, .21].

Discussion

As in the previous experiments, the data from Experiment 3 show that unitized associations do not generalize to facilitate learning of new related information. An additional effect that

emerged in this experiment is a main effect of relearning condition, with better memory following related versus unrelated relearning. This effect is likely due to the nature of the relearning task, in which participants are encouraged to bring to mind the definitions and sentences from the study phase when performing a memory judgment during relearning. Noticeably, such reference would be useful for memory judgments in repeated and related trials, where the information during relearning relates to that during study. This would facilitate information processing at relearning, and subsequently promote memory for these trials at test. Nevertheless, because this effect was equally strong for definition and sentence encoding, it does not entail any new insights regarding unitization.

Post-hoc analyses

Before turning to the *General discussion*, we describe a set of additional post-hoc analyses, meant to ensure that the null effect obtained in this study is indeed reliable. The first analysis was conducted to account for the fact that Bayes factors are sensitive to sample size. Because the structure of the three experiments was similar, we pooled the data from all three experiments (n = 64) to increase statistical power. As before, we took the difference in accuracy rates for definition versus sentence encoding, separately for related and unrelated relearning (see above), and used a one-sided Bayesian t-test with a Cauchy prior scaled at medium scaling, to compare the null hypothesis that the difference between encoding conditions does not differ for related and unrelated relearning, with the alternate hypothesis that this difference is bigger for related versus unrelated relearning. This analysis provided strong support for the null hypothesis over the alternate schema account, BF(h0) = 12.84, CrI [-.05 .02].

The second post-hoc analysis followed the notion that even a meta-analysis which combines only few results can give a valuable increase in precision (Cumming, Fidler, Kalinowski, & Lai, 2012). To estimate the pooled effect size of the Encoding X Relearning interaction across the three experiments, we calculated means and standard errors for the two difference measures (difference between definition and sentence encoding for related relearning; difference between definition and sentence encoding for unrelated relearning), and performed a Bayesian meta-analysis using the bmeta R package (Ding & Baio, 2016), which uses Markov Chain Monte Carlo (MCMC). In our implementation, the mean difference was the outcome measure, and the data were defined as continuous, two armed data. As shown in Fig. 6, the overall pooled μ was -0.016 (95% CrI [-0.14 0.111]), indicating that across all studies there was no indication for greater advantage of definition over sentence encoding conditions for related than for unrelated relearning conditions.
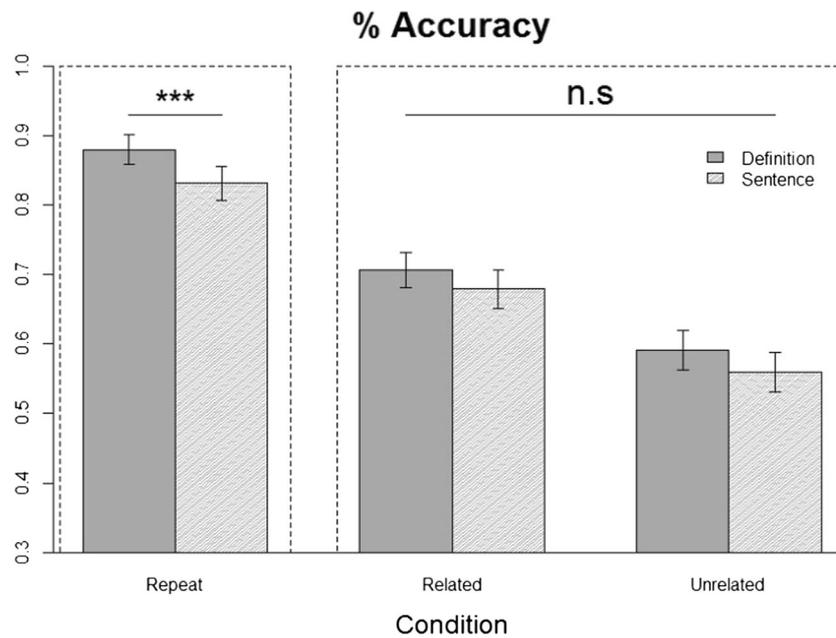
% Accuracy

Fig. 5 Experiment 3: Proportion of correct responses at test, in the various encoding and relearning conditions. The two left-most bars represent accuracy in the repeat (control) condition. Performance for repeated trials was significantly better for the definition than sentence conditions, consistent with the basic unitisation effect. The four bars on the right represent accuracy in the related and unrelated conditions (***p < .005). The predicted interaction between study condition and relearning condition was not significant (n.s). Error bars represent SEs for each condition separately

The aim of the third post-hoc analysis was to ensure that the study was sufficiently powered to detect potential effects. To this end, we calculated the observed power with the pwr R package (Champely, 2017) for the unitization effect obtained from Experiments 1 and 3 (n=48; Experiment 2 did not include the repeat conditions). This analysis revealed that the observed power was .99904, suggesting that the current study is very well powered for the detection of the unitization effect. Although this effect would arguably be the upper boundary of any generalization effects, the robustness of the unitization effect, coupled with a strong Bayesian preference of the null hypothesis, reinforce that the observed null effect is reliable.

The final post-hoc analysis aimed to rule out an alternative explanation – that this null effect is related to linguistic features of the to-be-learned compounds; more specifically, to their semantic transparency. Open compounds (i.e., compounds that are written as separate words) can differ in their level of transparency or opacity. Some compounds are fully transparent. In such compounds, the meaning of both constituents is transparently related to the meaning of the compound as a whole. In SCHOOL BUS, for example, the meaning of the compound word is related to the meaning of both constituents, as school bus is a vehicle (bus) that drives children to an educational institution (school). Turning to an example from the current study, under the definition of "a garden for skygazing", the compound CLOUD LAWN is fully transparent, as this a garden (lawn) from which one can gaze at the sky (clouds). In contrast, a compound like "cottage pie" is partially opaque, because, although it has to do with pie, it has less to do with cottage (semantically, rather than historically). Transparency level of the unitized compound might interact with its ability to generalize to facilitate new learning. Specifically, it could be that generalization is only possible when constituents are transparent, and therefore newly learned information can be linked back to the original

Table 5 Error data for Experiment 3. Data represent % of errors that made to each lure type. For each response type in each condition, mean and standard deviations are only calculated for participants who generated at least one erroneous response of this type

	Repeat			Related			Unrelated		
	Old-Different	New-Related	New-Unrelated	Old-Different	New-Related	New-Unrelated	Old-Different	New-Related	New-Unrelated
Definition	41.5 (32.8)	38.5 (28.9)	20.0 (21.5)	44.1 (14.9)	41.9 (15.3)	14.0 (10.4)	49.6 (13.6)	30.7 (12.1)	19.7 (8.9)
Sentence	38.7 (27.3)	38.4 (25.3)	22.9 (25.6)	43.5 (13.2)	34.3 (12.8)	22.2 (8.3)	52.2 (16.3)	25.3 (13.9)	22.5 (9.6)

Table 6 Proportion of new-related distractors in Experiment 3, for n=17 participants

	Repeat	Related	Unrelated
Definition	65.8 (29.9)	72.5 (19.3)	57.5 (16.4)
Sentence	66.1 (29.2)	59.2 (15.4)	53.2 (20.4)

meaning of the constituents. In the present study, in order to enable rapid processing of compounds that are learned within the context of single-trial learning, we used a stimulus database in which most compounds are fully transparent. Nevertheless, when validating the dataset we did not control for this property of the compounds. The inclusion of opaque compounds might have compromised our ability to obtain the predicted generalization effect. To rule out this potential alternative account, an individual (who did not participate in any experiment) reviewed the database, and marked all compounds that were not fully transparent. This resulted in 11 stimuli marked as partially transparent. All other stimuli were marked as fully transparent. We then excluded these entries from the dataset, pooled the data from all experiments (but only included trials where stimuli were fully transparent), and ran the same analysis which was used for our first post-hoc analysis. Once again, the analysis provided strong support for the null hypothesis over the alternate schema account, BF(h0) = 13.12, CrI [-.08 .03], suggesting that even when all the pairs are fully transparent, generalization does not occur.

General discussion

In three experiments, we tested an alternative account of unitization. We proposed that rather than fusing two items together to form a single unit (the traditional item account), unitization rapidly incorporates the items into a schema (the schema account). We employed a three-stage paradigm, in which unitized and non-unitized associations were learned in an initial study phase, re-paired with either related or unrelated information in a subsequent relearning phase, and retrieved in a final test phase. More specifically, we used definition versus sentence encoding during study to define unitization, and compared memory for related versus unrelated words presented in the relearning phase. If schemas enable generalization, then we predicted a greater advantage for related than unrelated words after definition encoding than after sentence encoding. After failing to find evidence for this interaction in Experiment 1, we conducted two further experiments to try to increase the chance of participants using schemas. In Experiment 2, the to-be-learned association was repeated twice at study with a different constituent to help indicate that it is generalizable, while in Experiment 3, we explicitly re-activated the information used at study during the relearning phase. Neither manipulation confirmed our prediction, and Bayes Factors favored the null hypothesis of no generalization. Thus, we found no evidence for the schema account of unitization, and conclude that the item account remains the most likely account to date.
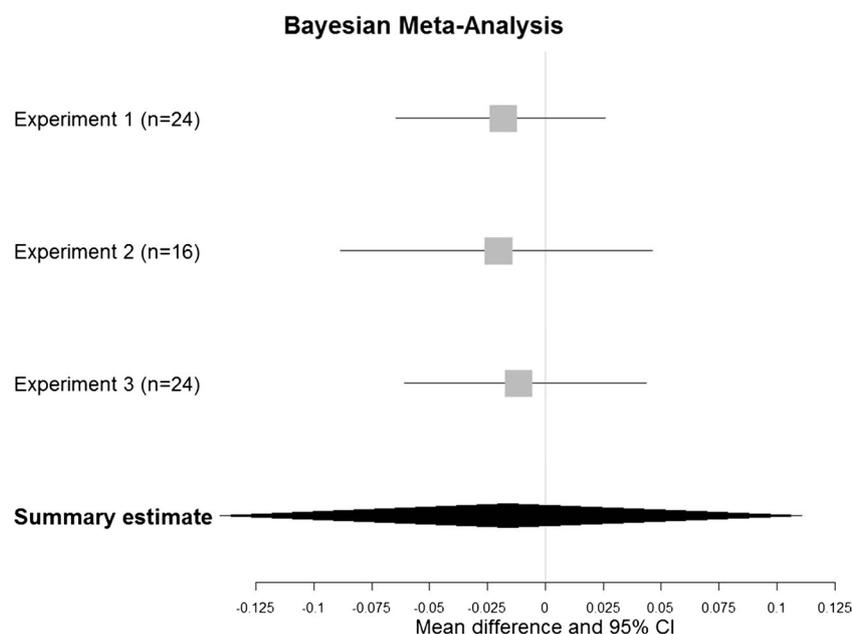

Bayesian Meta-Analysis

Fig. 6 Forest plot showing study-specific estimates of the mean difference for Experiments 1–3 (squares) and the pooled mean difference (diamond). On the x-axis, positive numbers indicate that an outcome favors the related condition (i.e., greater difference between definition and sentence encoding for related than for unrelated relearning), as predicted by the schema hypothesis, while negative numbers indicate an outcome that favors the unrelated condition. In the current case, the outcome was approximately 0, supporting the null hypothesis

We proposed that when a definition is presented at study, it establishes a schema. For example, even though a "COULD LAWN" was unlikely to be a concept already possessed by our participants, we hypothesized that defining it as "a garden used for sky-gazing" would provide a new structure that can accommodate both words, and subsequently would allow rapid one-shot learning of new, related associates. Is it possible that generalization of unitized associations does occur, but only when well-established, pre-existing schemas are involved? Some conceptualizations of schemas argue that they are only developed over time after repeated presentations, through extraction of commonalities across events (e.g., Ghosh & Gilboa, 2014). Furthermore, the acquisition of a schema, according to this definition, is believed to depend on systems consolidation (e.g., Tse et al., 2007), which would take longer than entailed by the current single-session study. Under these restrictions that schemas require multiple repetitions and/or consolidation, they would be unable to ever explain unitization effects, since these occur under single-trial, same-day learning. In the current study, we used a modified definition of schema, based on the SLIMM model (van Kesteren et al., 2012). In this framework, the term "schema" refers to a cognitive construct with a specific structure (a network of interconnected and activated representations) and function (affecting online information processing, including the acquisition of new knowledge). We provide evidence that such construct (whether termed "schema" or otherwise referred to as "script" or "gist") does not underlie unitization.

If abstract structures like schemas can be established quickly, into which related concepts can be assimilated, then it is possible that they were generated in our sentence condition too (but simply less effectively than in our definition condition, given that memory was better after definition than sentence encoding in the repeat condition). If schemas were established that allow generalization in both encoding conditions, then this might explain why there was minimal interaction between the encoding conditions and related versus unrelated relearning. However, if schemas allow generalization, then we would expect a main effect of relearning condition, with better memory for related than unrelated words under both encoding conditions. This main effect was seen in Experiment 3, but not in Experiments 1 and 2. Thus, any evidence for generalization – the defining feature that distinguishes the schema account of unitization from the item account – remains limited.

Another possibility that should be considered is that the generalization effect was not obtained because the tasks that were used during the relearning phase were not sufficient to impact performance at the test phase. Indeed, in Experiments 1 and 2, participants were not explicitly instructed to remember the items during the relearning phase, and thus encoding may have been incidental. Nevertheless, we believe that substantial learning did occur during this phase for two reasons.

First, although participants were not instructed to remember the stimuli during relearning, they knew that a test phase would follow, in which they would have to retrieve this information. This is because they were informed about the structure of the task, and also completed a practice block which included all three phases. Second, because the target stimuli in the unrelated conditions were semantically unrelated to the stimuli from the initial study phase, above-chance accuracy rates in this condition can only reflect learning that occurred during the relearning phase. The fact that accuracy rates in these conditions were well above chance in all experiments indicates that substantial learning did occur during the relearning phase.

Although we did not find any evidence to support a schema account of unitization, the converse may nevertheless be the case, i.e., that some of the schema effects reported in the literature actually reflect unitization. For example, in a study aiming to identify the neural correlates of schema-dependent encoding, van Kesteren, Beul, Takashima, Henson, Ruiter and Fernández (2013) presented scene-item pairs at varying congruency levels; more recently, Liu, Grady, and Moscovitch (2016) showed participants houses that were paired with faces of either famous or non-famous people, to examine effects of prior knowledge (as a simplified form of schema) on memory encoding; Packard, Rodríguez-Fornells, Bunzeck, Nicolás, de Diego-Balaguer and Fuentemilla (2017; Experiment 4) used category cues paired with either a congruent or incongruent target word to examine the temporal dynamics of schema effects on memory. In all of these examples, the two encoded items may have been bound together into a single item, and more easily so in the congruent than the incongruent conditions, and therefore the results of these studies can potentially/partially be explained by unitization. Future research investigating schema effects on memory can benefit from our current study by more carefully examining the properties of the learned associations: if they generalize to accommodate new information, than they most likely reflect a schema rather than a unit.

Although the definition/sentence procedure used here is a common unitization manipulation, other unitization manipulations have been used in previous studies. Unitization can be achieved by two broad categories of experimental manipulations, driven by either top-down or bottom-up cognitive processes (Tibon et al., 2014). Top-down unitization, which was the focus of the current investigation, is driven purely by instructions to process pairs of unrelated memoranda as a single unit (in high unitization conditions) or as separate elements of the same episode (in low unitization conditions). Bottom-up unitization leverages pre-existing semantic and/or perceptual relations between to-be-associated memoranda. Especially relevant to a schema account are bottom-up manipulations that rely on pre-existing semantic links between the encoded items. For example, existing word compounds such as

"cottage pie" are already unitized prior to an experiment, and are therefore more likely to be perceived as a unit during the experiment than new word pairs (e.g., Giovanello, Keane, & Verfaellie, 2006; Rhodes & Donaldson, 2007; and see also Tibon et al, 2014 for similar results with visual stimuli). Arguably, in such cases, a schema account might be more applicable than in the case of top-down unitization. Namely, it is possible that when bottom-up unitization is employed, reinforcing existing links would be a more efficient process than fusing items together to create a new representation. To the extent that different mechanisms support unitization under different circumstances, unitization might be more likely to act as a schema in bottom-up situations. Nevertheless, our current data speak only to top-down unitization, so this would be an interesting prediction to investigate in future studies.

Additional consideration should be given to other top-down unitization manipulations. These include viewing or imagining two stimuli as interacting versus independent (D'Angelo et al., 2017; Rhodes & Donaldson, 2008), or encoding source and item information in an internal versus external manner, thus forming intra- versus inter-item associations (e.g., "imagine the item in the color indicated by the background screen color" in the unitization condition, versus "imagine why the item would be associated with a stop sign or dollar bill" in the non-unitization condition; Bastin et al., 2013; Diana et al., 2011). It is possible that our claim to rule out a schema account is limited to the manipulation employed in the current study, and it may not apply to other top-down manipulations. Nonetheless, we think the definition manipulation is the one that is most likely to produce a schema, because it explicitly entails the activation of an abstract concept. For other top-down manipulations, such as visualizing an item in a color indicated by the frame, a direct link can be established between the two stimuli (in this case, an item and a color), making a schema account less likely.

In summary, Bayesian analysis of the data in the current study provides strong evidence against a key prediction of the schema account of unitization, namely that schema should generalize to related information. While the default item account therefore remains a likely explanation, alternative explanations of unitization should be considered in future studies. In particular, studies need to go beyond the finding that unitized associations can be retrieved via familiarity, by also verifying other item-like properties that are enabled by unitization.

References

Ahmad, F. N., & Hockley, W. E. (2017). Distinguishing familiarity from fluency for the compound word pair effect in associative recognition. *The Quarterly Journal of Experimental Psychology, 70*(9), 1768–1791. https://doi.org/10.1080/17470218.2016.1205110

Anderson, J. R. (1981). Effects of prior knowledge on memory for new information. *Memory & Cognition*, 9(3), 237–246. https://doi.org/10.3758/BF03196958

Bader, R., Mecklinger, A., Hoppstädter, M., & Meyer, P. (2010). Recognition memory for one-trial-unitized word pairs: Evidence from event-related potentials. *NeuroImage*, 50(2), 772–781. https://doi.org/10.1016/j.neuroimage.2009.12.100

Bader, R., Opitz, B., Reith, W., & Mecklinger, A. (2014). Is a novel conceptual unit more than the sum of its parts?: FMRI evidence from an associative recognition memory study. *Neuropsychologia, 61*, 123–134. DOI: https://doi.org/10.1016/j.neuropsychologia.2014.06.006

Bartlett, F.C., (1932). Remembering: An Experimental and Social Study. Cambridge Univ, Cambridge.

Bastin, C., Diana, R. A., Simon, J., Collette, F., Yonelinas, A. P., & Salmon, E. (2013). Associative memory in aging: The effect of unitization on source memory. *Psychology and Aging*, 28(1), 275. https://doi.org/10.1037/a0031566

Champely, S. (2017). pwr: Basic functions for power analysis. R package version 1.2–1. https://CRAN.R-project.org/package=pwr

Craik, F. I., & Bosman, E. A. (1992). Age-related changes in memory and learning. *Gerontechnology*, 3, 79.

Cumming, G., Fidler, F., Kalinowski, P., & Lai, J. (2012). The statistical recommendations of the American psychological association publication manual: Effect sizes, confidence intervals, and metaanalysis. *Australian Journal of Psychology, 64*(3), 138–146. https://doi.org/10.1111/j.1742-9536.2011.00037.x

D'Angelo, M. C., Kacollja, A., Rabin, J. S., Rosenbaum, R. S., & Ryan, J. D. (2015). Unitization supports lasting performance and generalization on a relational memory task: Evidence from a previously undocumented developmental amnesic case. *Neuropsychologia*, 77, 185–200. https://doi.org/10.1016/j.neuropsychologia.2015.07.025

D'Angelo, M. C., Noly-Gandon, A., Kacollja, A., Barense, M. D., & Ryan, J. D. (2017). Breaking down unitization: Is the whole greater than the sum of its parts?. *Memory & cognition*, 45(8), 1306–1318

Delhaye, E., & Bastin, C. (2016). The impact of aging on associative memory for preexisting unitized associations. *Aging, Neuropsychology, and Cognition*, 1–29. https://doi.org/10.1080/13825585.2016.1263725

Delhaye, E., Tibon, R., Gronau, N., Levy, D. A., & Bastin, C. (2017). Misrecollection prevents older adults from benefitting from semantic relatedness of the memoranda in associative memory. *Aging, Neuropsychology, and Cognition*, 1–21. https://doi.org/10.1080/13825585.2017.1358351

Diana, R. A., Van den Boom, W., Yonelinas, A. P., & Ranganath, C. (2011). ERP correlates of source memory: Unitized source information increases familiarity-based retrieval. *Brain Research*, 1367, 278–286. DOI: https://doi.org/10.1016/j.brainres.2010.10.030

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a

behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 730. https://doi.org/10.1037/0278-7393.34.4.730

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2010). Medial temporal lobe activity during source retrieval reflects information type, not memory strength. *Journal of Cognitive Neuroscience, 22*(8), 1808–1818. https://doi.org/10.1162/jocn.2009.21335

Ding, T., & Baio, G. (2016). bmeta: Bayesian meta-analysis and meta-regression. R package version 0.1.2. https://CRAN.R-project.org/package=bmeta

Donaldson, D. I., & Rugg, M. D. (1998). Recognition memory for new associations: Electrophysiological evidence for the role of recollection. *Neuropsychologia, 36*(5), 377–395. https://doi.org/10.1016/S0028-3932(97)00143-7

Ford, J. H., Verfaellie, M., & Giovanello, K. S. (2010). Neural correlates of familiarity-based associative retrieval. *Neuropsychologia, 48*(10), 3019-3025.

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia, 53*, 104–114. https://doi.org/10.1016/j.neuropsychologia.2013.11.010

Giovanello, K. S., Keane, M. M., & Verfaellie, M. (2006). The contribution of familiarity to associative memory in amnesia. *Neuropsychologia, 44*(10),1859–1865. https://doi.org/10.1016/j.neuropsychologia.2006.03.004

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5*(6), 236–243. https://doi.org/10.1016/S1364-6613(00)01662-4

Graf, P., & Schacter, D. L. (1989). Unitization and grouping mediate dissociations in memory for new associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(5), 930. https://doi.org/10.1037/0278-7393.15.5.930

Guillaume, F., & Etienne, Y. (2015). Target-context unitization effect on the familiarity-related FN400: A face recognition exclusion task. *International Journal of Psychophysiology, 95*(3), 345–354. https://doi.org/10.1016/j.ijpsycho.2015.01.004

Haskins, A. L., Yonelinas, A. P., Quamme, J. R., & Ranganath, C. (2008). Perirhinal cortex supports encoding and familiarity-based recognition of novel associations. *Neuron, 59*(4), 554–560. https://doi.org/10.1016/j.neuron.2008.07.035

Hockley, W. E., & Consoli, A. (1999). Familiarity and recollection in item and associative recognition. *Memory & Cognition, 27*(4), 657–664. https://doi.org/10.3758/BF03211559

Jäger, T., Mecklinger, A., & Kipp, K. H. (2006). Intra-and inter-item associations doubly dissociate the electrophysiological correlates of familiarity and recollection. *Neuron, 52*(3), 535–545. https://doi.org/10.1016/j.neuron.2006.09.013

Jäger, T., Mecklinger, A., & Kliegel, M. (2010). Associative recognition memory for faces: More pronounced age-related impairments in binding intra-than inter-item associations. *Experimental Aging Research, 36*(2), 123–139. https://doi.org/10.1080/03610731003613391

Kamp, S. M., Bader, R., & Mecklinger, A. (2016). The Effect of Unitizing Word Pairs on Recollection Versus Familiarity-Based Retrieval—Further Evidence From ERPs. *Advances in Cognitive Psychology, 12*(4), 169. https://doi.org/10.5709/acp-0196-2

Kikyo, H., Ohki, K., & Miyashita, Y. (2002). Neural correlates for feeling-of-knowing: An fMRI parametric analysis. *Neuron, 36*(1), 177–186. https://doi.org/10.1016/S0896-6273(02)00939-X

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Lawrence, A. (2016). ez: Easy analysis and visualization of factorial experiments. R package version 4.4–0. https://CRAN.R-project.org/package=ez

Liu, Z. X., Grady, C., & Moscovitch, M. (2016). Effects of prior-knowledge on brain activation and connectivity during associative memory encoding. *Cerebral Cortex, 27*(3), 1991–2009. https://doi.org/10.1093/cercor/bhw047

Mayes, A., Montaldi, D., & Migo, E. (2007). Associative memory and the medial temporal lobes. *Trends in Cognitive Sciences, 11*(3), 126–135. https://doi.org/10.1016/j.tics.2006.12.003

Mecklinger, A. (2000). Interfacing mind and brain: A neurocognitive model of recognition memory. *Psychophysiology, 37*(5), 565–582. https://doi.org/10.1017/S0048577200992230

Mecklinger, A., & Jäger, T. (2009). Episodic memory storage and retrieval: Insights from electrophysiological measures. In F. Rösler, C. Ranganath, B. Röder, R.H. Kluwe (Eds.), *Neuroimaging and psychological theories of human memory, 357–382. Oxford: Oxford University Press.

Memel, M., & Ryan, L. (2017). Visual integration enhances associative memory equally for young and older adults without reducing hippocampal encoding activation. *Neuropsychologia, 100*, 195–206. https://doi.org/10.1016/j.neuropsychologia.2017.04.031

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: computation of bayes factors for common designs. R package version 0.9.12–2. https://CRAN.R-project.org/package=BayesFactor

Packard, P. A., Rodríguez-Fornells, A., Bunzeck, N., Nicolás, B., de Diego-Balaguer, R., & Fuentemilla, L. (2017). Semantic congruence accelerates the onset of the neural signals of successful memory encoding. *Journal of Neuroscience, 37*(2), 291–301. https://doi.org/10.1523/JNEUROSCI.1622-16.2016

Paller, K. A., Voss, J. L., & Boehm, S. G. (2007). Validating neural correlates of familiarity. *Trends in Cognitive Sciences, 11*(6), 243–250. https://doi.org/10.1016/j.tics.2007.04.002

Parks, C. M., & Yonelinas, A. P. (2015). The importance of unitization for familiarity-based learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(3), 881. https://doi.org/10.1037/xlm0000068

Pilgrim, L. K., Murray, J. G., & Donaldson, D. I. (2012). Characterizing episodic memory retrieval: Electrophysiological evidence for diminished familiarity following unitization. *Journal of Cognitive Neuroscience, 24*(8), 1671–1681. https://doi.org/10.1162/jocn_a_00186

Quamme, J. R., Yonelinas, A. P., & Norman, K. A. (2007). Effect of unitization on associative recognition in amnesia. *Hippocampus, 17*(3), 192–200. https://doi.org/10.1002/hipo.20257

Revelle, W. (2017) psych: Procedures for personality and psychological research, Northwestern University, Evanston, Illinois, USA. R package version 1.7.8. https://CRAN.R-project.org/package=psych

Rhodes, S. M., & Donaldson, D. I. (2007). Electrophysiological evidence for the influence of unitization on the processes engaged during episodic retrieval: Enhancing familiarity based remembering. *Neuropsychologia, 45*(2), 412–424. https://doi.org/10.1016/j.neuropsychologia.2006.06.022

Rhodes, S. M., & Donaldson, D. I. (2008). Electrophysiological evidence for the effect of interactive imagery on episodic memory: Encouraging familiarity for non-unitized stimuli during associative recognition. *NeuroImage, 39*(2), 873–884. https://doi.org/10.1016/j.neuroimage.2007.08.041

Robey, A., & Riggins, T. (2017). Increasing relational memory in childhood with unitization strategies. *Memory & Cognition, 1–12. https://doi.org/10.3758/s13421-017-0748-6

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257. https://doi.org/10.1016/j.tics.2007.04.004

Rumelhart, D. E. (1980). On evaluating story grammars. *Cognitive Science*, *4*(3), 313–316. https://doi.org/10.1016/S0364-0213(80)80007-3

Ryan, J. D., Moses, S. N., Barense, M., & Rosenbaum, R. S. (2013). Intact learning of new relations in amnesia as achieved through unitization. *Journal of Neuroscience*, *33*(23),9601–9613. https://doi.org/10.1523/JNEUROSCI.0169-13.2013

Shao, H., Opitz, B., Yang, J., & Weng, X. (2016). Recollection reduces unitised familiarity effect. *Memory*, *24*(4), 535–547. https://doi.org/10.1080/09658211.2015.1021258

Tibon, R., Ben-Zvi, S., & Levy, D. A. (2014). Associative recognition processes are modulated by modality relations. *Journal of Cognitive Neuroscience*, *26*(8), 1785–1796. https://doi.org/10.1162/jocn_a_00586

Tibon, R., Gronau, N., Scheuplein, A. L., Mecklinger, A., & Levy, D. A. (2014). Associative recognition processes are modulated by the semantic unitizability of memoranda. *Brain and Cognition*, *92*, 19–31. https://doi.org/10.1016/j.bandc.2014.09.009

Tibon, R., Vakil, E., Goldstein, A., & Levy, D. A. (2012). Unitization and temporality in associative memory: Evidence from modulation of context effects. *Journal of Memory and Language*, *67*(1), 93–105. https://doi.org/10.1016/j.jml.2012.02.003

Troyer, A. K., D'Souza, N. A., Vandermorris, S., & Murphy, K. J. (2011). Age-related differences in associative memory depend on the types of associations that are formed. *Aging, Neuropsychology, and Cognition*, *18*(3), 340–352. https://doi.org/10.1080/13825585.2011.553273

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., … Morris, R. G. (2007). Schemas and memory consolidation. *Science*, *316*(5821), 76–82. https://doi.org/10.1126/science.1135935

Tu, H. W., Alty, E. E., & Diana, R. A. (2017). Event-related potentials during encoding: Comparing unitization to relational processing. *Brain Research*, *1667*, 46–54. https://doi.org/10.1016/j.brainres.2017.05.003

Tulving, E. (1972). Episodic and semantic memory. *Organization of memory*, *1*, 381–403.

van Kesteren, M. T., Beul, S. F., Takashima, A., Henson, R. N., Ruiter, D. J., & Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. *Neuropsychologia*, 51(12), 2352–2359. https://doi.org/10.1016/j.neuropsychologia.2013.05.027

van Kesteren, M. T., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211–219. https://doi.org/10.1016/j.tins.2012.02.001

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, *277*(5324), 376–380. https://doi.org/10.1126/science.277.5324.376

Wilding, E. L., & Ranganath, C. (2011). Electrophysiological correlates of episodic memory processes. *The Oxford handbook of ERP components*, 373–396.

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, *25*(6), 747–763. https://doi.org/10.3758/BF03211318

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194. https://doi.org/10.1002/hipo.20864

Yonelinas, A. P., Kroll, N. E. A., Dobbins, I. G., & Soltani, M. (1999). Recognition memory for faces: When familiarity supports associative recognition judgments. *Psychonomic Bulletin & Review*, *6*(4), 654–661. https://doi.org/10.3758/BF03212975

Zheng, Z., Li, J., Xiao, F., Broster, L. S., Jiang, Y., & Xi, M. (2015). The effects of unitization on the contribution of familiarity and recollection processes to associative recognition memory: Evidence from event-related potentials. *International Journal of Psychophysiology*, 95(3), 355–362. https://doi.org/10.1016/j.ijpsycho.2015.01.003