



**HAL**  
open science

## Children's Gist-based False Memory in Working Memory Tasks

Manon Rousselle, Marlène Abadie, Agnès Blaye, Valérie Camos

► **To cite this version:**

Manon Rousselle, Marlène Abadie, Agnès Blaye, Valérie Camos. Children's Gist-based False Memory in Working Memory Tasks. *Developmental Psychology*, In press. hal-03779489

**HAL Id: hal-03779489**

**<https://hal.science/hal-03779489>**

Submitted on 16 Sep 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Running title: False memories in children's WM

In press, 2022, *Developmental Psychology*

This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission. The final article will be available, upon publication.

## Children's Gist-based False Memory in Working Memory Tasks.

Manon Rousselle<sup>1</sup>, Marlène Abadie<sup>1</sup>, Agnès Blaye<sup>1</sup> & Valérie Camos<sup>2</sup>

### Author note

1 Manon Rousselle, Marlène Abadie & Agnès Blaye, Aix Marseille Université, CNRS, LPC, Marseille, France. 2 Valérie Camos, Département de Psychologie, Université de Fribourg, Suisse. This work was supported by a grant from the French National Research Agency (No. ANR-19-CE28-0017-01) to Marlène Abadie. The funding source was not involved in the conduct of the research nor in the preparation of the article. Data and material for this study are openly available at the Open Science Framework (OSF), <https://osf.io/9gfmd/>. We thank Michaël Chabaud and Camille Watelain for their assistance with data collection and Christelle Guette for proofreading this manuscript. This study was not preregistered. Correspondence concerning this article should be addressed to Manon Rousselle or Marlène Abadie, CNRS & AMU, UMR7290, Bâtiment 9 Case D, 3, place Victor Hugo, 13331 Marseille Cedex 3, France. manon.rousselle-jacquier@univ-amu.fr or marlene.abadie@univ-amu.fr, phone: +33(0) 4 13 55 09 60 or +33(0) 4 13 55 09 74.

## Abstract

False memories are well established episodic memory phenomena. Recent research in young adults has shown that semantically related associates can be falsely remembered as studied items in working memory (WM) tasks for lists of only a few items when a short 4-second interval was given between study and test. The present study reported two experiments yielding similar effects in 4- ( $n = 32$  and  $33$ ,  $18$  and  $14$  females, respectively) and 8-year-old children ( $n = 33$  and  $34$ , respectively,  $19$  females in both). Short lists of semantically related items specifically tailored for young children were retained over a brief interval. Whether or not the interval was filled with a concurrent task that impeded or not WM maintenance, younger children were as prone to falsely recognize related distractors as their older counterparts in an immediate recognition test, and also in a delayed test. In addition, using the conjoint recognition model of the fuzzy-trace theory, we demonstrated that the retrieval of gist traces of the list themes was responsible for the occurrence of short-term false memories in 4- and 8-year-old children. Gist memory also underpinned the occurrence of false recognition in the delayed test. These findings suggest that young children are as likely to make gist-based false memories as older children in working memory tasks.

Keywords: False Memory; Cognitive Development; Fuzzy Trace Theory; Working Memory; Long-term Memory

Research on false memories in children has been the subject of intense debate over the past 30 years. It emerged from concerns about false reports by young children in certain types of legal cases (e.g., *Buckey v. County of Los Angeles*, 1992). The question that has received much attention is whether young children are more vulnerable to false memories than older children or adults. Research has first shown that children's memories, especially of younger children, are easily distorted by suggestions, whether they are blatant and deliberate or subtle and involuntary (for a review, see Howe, 2000). In addition to these early studies, much evidence has accumulated showing that young children are generally less susceptible to the well-known Deese-Roediger-McDermott illusion (DRM, Deese, 1959; Roediger & McDermott, 1995) than older children and adults (e.g., Anastasi & Rhodes, 2008; Brainerd et al., 2018). False memories in the DRM paradigm are well-established episodic long-term memory (LTM) phenomena: within minutes or days of learning lists composed of many semantic associates, other unstudied semantic associates are confidently and erroneously remembered as studied items. This powerful illusion does not depend on providing children with suggestive materials or leading questions, but is rather the consequence of the development of associative processes that enables older children to make meaning-based inferences about experienced events. Recent studies in young adults have shown that, surprisingly, the DRM illusion could also occur in working memory (WM) tasks for lists of only a few items when a short 4-second interval was given between study and test (e.g., Abadie & Camos, 2019; Atkins & Reuter-Lorenz, 2008). However, this illusion in WM has never been studied in children. This question is of particular interest since WM is not fully developed in children (e.g., Camos & Barrouillet, 2018; Cowan, 2014). The aims of the present study were, first, to examine whether false memories could also occur in WM tasks in young children and whether they are prone to the same age-related increase as long-term false

memories in the DRM paradigm and, second, to determine the role of WM in their occurrence.

### **The development of long-term false memories**

In the DRM paradigm, participants study lists composed of the 15 strongest associates (e.g., “climber, peak, ski, valley, etc.”) to a given theme word (e.g., “mountain”) that is not itself presented, and are then given a recall or a recognition test that may occur minutes or days later. This paradigm typically produces high levels of false recall or recognition of the critical theme word in young adults, often rivaling the acceptance rate for presented words (e.g., Reyna & Lloyd, 1997; Roediger & McDermott, 1995). According to the Fuzzy-Trace theory (FTT, Brainerd & Reyna, 2002a; Brainerd et al., 2008; Chang & Brainerd, 2021), subjects store two types of episodic representations as they encode the items in DRM lists: verbatim traces of item presentations and gist traces of item meanings, especially semantic relations that connect the different items of a given list. Verbatim traces are representations of surface features of items, i.e., auditory, visual, or other details associated with the item presentation (e.g., the color, font, size, position, etc. of an item; Brainerd et al., 2019). Gist traces are representations of the semantic content of items and other relational information such as taxonomic, synonymous, or situational relations. Access to verbatim and gist memories declines over time, but verbatim traces are faster to fade away (Brainerd & Reyna, 2002b) and are also more fragile and sensitive to cognitive load and interferences than gist traces (Abadie & Waroquier, 2020; Abadie et al., 2013; 2017). The FTT states that the retrieval of verbatim traces supports true memory for list words and suppresses false memories for critical distractors (remembering “climber”, “peak”, produces rejection of “mountain”; Brainerd et al., 2003), while gist retrieval supports both true and false memories (remembering that the gist of the list was “mountain” produces acceptance of “climber” but also of “mountain”; Brainerd & Reyna, 2005). Because they are associated,

DRM-list words repeatedly cue list themes (e.g., “mountain”) resulting in very strong gist memory of these themes, which are especially likely to be falsely recalled or recognized.

Many studies have investigated the occurrence of the DRM illusion in children in relation to the development of verbatim and gist traces. Brainerd and Reyna (1998) and Ceci and Bruck (1998) were the first to propose that false memories increase with age. Soon, several investigators found that the false memory effect is at near floor levels in young children aged 5 to 7 years and intensifies throughout childhood and adolescence (Anastasi & Rhodes, 2008; Brainerd et al., 2018; Sugrue & Hayne, 2006; Wimmer & Howe, 2009). The FTT accounts for these findings by stating that acquisition, retention and retrieval of both gist and verbatim memories improve from childhood to adulthood (see Brainerd & Reyna, 2002a; Brainerd & Reyna, 2015, for reviews). Although both verbatim and gist memories increase from early childhood onwards, the increase is sharper for gist than for verbatim traces (Brainerd et al., 2002c; Reyna & Kiernan, 1994). Gist memory improves during childhood, because the ability to process meaning of an item and to connect meaning across different items improves. Semantic clustering in long-term recall or recognition increases during childhood, and intrusions become increasingly semantic with age. It is well established that the magnitude of this trend can be influenced by many theoretically driven manipulations. For instance, encouraging 7- and 11-year-olds to search for meaningful relations when items of DRM lists are encoded or when DRM lists are repeatedly tested boosts reliance on gist, leading to an increase in long-term false recognition for both age groups. By contrast, focusing on processing superficial characteristics of the stimuli, such as spelling, reduces false memories in both age groups (Holliday et al., 2011; Odegard et al., 2008). Other studies presented DRM-list words in stories that revolved around the lists’ themes (Dewhurst et al., 2007; Howe & Wilkinson, 2011). Stories increased long-term false memories in 5-year-olds but had little effect in 8- and 11-year-olds.

## **Toward a new account of short- and long-term false memories**

Although the phenomenon of false memories has been widely studied in the field of LTM, a few studies in young adults have reported that false memories can also occur in WM tasks within seconds of studying a small number of semantically associated words (e.g., Atkins & Reuter-Lorenz, 2008; Coane et al., 2007; Flegal et al., 2010; Flegal & Reuter-Lorenz, 2014). At first glance, these findings suggest that processes that give rise to false memories are not exclusive to LTM, but may also be active in WM, a limited capacity system for maintaining relevant representations over the short term. To thoroughly examine the role of WM on the false memory effect, a series of experiments were conducted in young adults using a Brown Peterson paradigm (Brown, 1958; Peterson & Peterson, 1959) with an immediate and a delayed recognition test of DRM-like lists (Abadie & Camos, 2019). Participants had to maintain four highly or weakly semantically related words for 4 seconds before recognizing them among related and unrelated distractors. The 4-second retention interval was filled by a concurrent task of varying difficulty to manipulate the possibility that the items could be actively maintained in WM. Two experiments demonstrated the role of WM in the occurrence of false memories: active maintenance of word-lists in WM prevented false recognition of related distractors in the immediate test but elicited false recognition in the delayed test. Abadie and Camos (2019) therefore proposed that active maintenance in WM would prevent the occurrence of false memories in short-term tests. It would also participate in their formation in long-term tests such as in the classical DRM paradigm.

In the time-based resource-sharing model (TBRM, Barrouillet & Camos, 2015), two mechanisms have been described as responsible for the maintenance of information in WM: articulatory rehearsal, which operates by articulatory repetition of memory items, and attentional refreshing, a domain-general attention-based mechanism that operates by briefly thinking back to recently active memory items (Camos, 2015, 2017). Studies demonstrated

that articulatory rehearsal and attentional refreshing had independent and additive effects on immediate recall. Manipulating the availability of one mechanism while controlling for the other led to poorer immediate recall performance. To disentangle the role of each maintenance mechanism on the emergence of false memories, Abadie and Camos (2019) varied the attentional demand of the concurrent task and the presence (or not) of a concurrent articulation during the retention interval of a Brown-Peterson task to impede either refreshing or rehearsal, respectively. Results revealed that using rehearsal prevents false recognition of related distractors in the immediate test whereas refreshing increases false recognition of related distractors in the delayed test.

By integrating the functional description of WM from the TBRS model with the FTT, Abadie and Camos (2019) proposed a new model that allows for a precise depiction of how each WM maintenance mechanism affects the creation of verbatim and gist traces and thus the formation of false memories. Consistent with the predictions of the FTT, this model predicts that false memories in both the immediate and delayed tests are underpinned by the retrieval of gist traces. Specifically, the use of rehearsal should promote the retrieval of verbatim traces that oppose gist traces in the immediate test, hence reducing the occurrence of short-term false memories. By contrast, the use of refreshing is assumed to foster the creation of both verbatim and gist traces. Gist traces are counteracted by verbatim traces in immediate tests, but as they are less sensitive to loss than the latter (e.g., Brainerd & Reyna, 2005), they can support memory retrieval after long delay of retention. Thus, impairing refreshing should reduce both verbatim and gist retrieval, resulting in fewer correct recalls or recognitions in immediate tests. However, in delayed tests, as performance relies more on gist retrieval, preventing the use of refreshing should decrease both true and false memories. These predictions on correct and false recognition as well as on measures of verbatim and gist representations were verified (Abadie & Camos, 2019).



Extending this model to childhood, a seemingly self-evident prediction is that the occurrence of false memories in a WM task will increase with age as it does in the classical DRM episodic memory task due to the development of gist memory. However, WM capacity is characterized by a significant age-related increase during childhood (Camos & Barrouillet, 2018; Cowan, 2014, for reviews). Memory span, the number of items successfully maintained in WM, increases throughout childhood, until early adolescence (e.g., Gathercole et al., 2004). Moreover, research showed that children of about 4-years of age and younger do not spontaneously attempt to refresh or rehearse memory items when engaged in a concurrent task. WM maintenance mechanisms become efficient from age 7 onwards and keep increasing in efficiency until adolescence, which underpins a sizable part of WM increase with age (Henry, 2012, for a review). Thus, if, as predicted by Abadie and Camos' (2019) model, WM plays an important role in preventing short-term false memories, young children under the age of 7 may not be able to maintain information correctly and may produce semantic errors in WM tasks. Regarding the use of each main WM maintenance mechanism, there is substantial evidence that rehearsal starts to be efficient between the ages of 5 and 7 (Henry et al., 2012; Tam et al., 2010), while children would start to use refreshing at age 7 (Barrouillet et al., 2009; Camos & Barrouillet, 2011). In addition, since the use of refreshing is attention-demanding, children favor the use of rehearsal whenever they can. Recent studies suggested that by the age of 6, children use rehearsal as a default maintenance strategy, while from age 8 onwards they could adaptively switch between rehearsal and refreshing (Oftinger & Camos, 2018). So far, the literature on WM development has focused on true memory, and no studies have yet investigated the role of these mechanisms in the formation of false memories in children.

### **The present study**

We report two experiments<sup>1</sup> that were designed to examine, first, whether false memories can occur in a WM task in children and whether they increase with age as in the classical DRM paradigm. Secondly, these experiments tested Abadie and Camos' (2019) model regarding the role of WM maintenance mechanisms in the occurrence of false memories and the underlying verbatim and gist representations. We compared false memories in 4- and 8-year-old children since the shift in the use of WM maintenance mechanisms happens between these ages. Therefore, 4-year-olds should not yet be using WM maintenance mechanisms, whereas 8-year-olds should spontaneously use maintenance strategies. Children were presented with DRM-like lists of 3-5 items represented by pictures and/or auditorily-presented words. The number of memory items was determined based on previous studies in children of similar ages (e.g., Barrouillet et al., 2009; Hitch et al., 1989). The presentation of each list was followed by a retention interval of a few seconds filled with a concurrent task in a Brown-Peterson paradigm. The type of concurrent task varied in each experiment to manipulate the possibility of implementing WM maintenance mechanisms. As in previous studies in adults, following the retention interval, an immediate recognition test composed of target items, related and unrelated distractors was proposed to the children. In Experiment 2, following the completion of all trials, participants completed a delayed recognition test similar to the immediate test. This test was added in order to compare the results with those obtained in the classical DRM paradigm. Finally, the simplified version of the conjoint recognition (SCR) model of the FTT (Brainerd, et al., 1999; Stahl & Klauer, 2008) was used to directly measure the verbatim and gist representations that underlie true and false memory in the immediate and delayed test. This model delivers precise quantitative measurements of the contributions of verbatim and gist memory for the responses to the different types of probes.

---

<sup>1</sup> A third experiment is reported as a supplementary material on the OSF. Its objectives and results are outlined in the discussion of Experiment 2.

Based on findings on the development of verbatim and gist memory (e.g., Brainerd et al., 2018), we expected that 8-year-old children would retrieve more verbatim and gist traces than 4-year-olds, which would increase both true recognition of target probes and false recognition of related distractors in the immediate and delayed tests in 8-year-olds. Because WM maintenance mechanisms are efficient from 7 (e.g., Barrouillet et al., 2009; Camos & Barrouillet, 2011), correct recognition in the immediate test should be reduced in 8-year-olds when the use of WM maintenance mechanisms was prevented relative to when it was not, whereas such a manipulation should not affect correct recognition in 4-year-olds. Most importantly, results in young adults showed that false recognition increased in the immediate test when the use of WM maintenance, specifically rehearsal, was blocked (Abadie & Camos, 2019). Therefore, we also expected that the rate of false recognition of related distractors would increase in 8-year-olds when they could not use any WM maintenance mechanism. In addition, as predicted by Abadie and Camos (2019), verbatim and gist traces retrieval should be affected by the manipulation of WM maintenance mechanisms, but only in 8-year-olds. For this age group, verbatim retrieval in the immediate test should be reduced when WM maintenance was prevented, while gist retrieval should increase. Finally, the present study also examined whether the use of WM maintenance strategies affects long-term information retention in young children. Studies indicated that the use of refreshing increased long-term true and false memories in young adults (Abadie & Camos, 2019; Camos & Portrat, 2015). However, no study has yet examined the long-term effects of using refreshing in children. Therefore, the delayed test in Experiment 2 aimed at examining this issue.

### **Experiment 1**

Four- and 8-year-old children performed a Brown-Peterson task in which they had to maintain lists of auditory-presented words, each of them being associated with a drawing. The words in a given list were all related to one another and to a common theme, just as in the

DRM lists, with the only difference being that relevant themes and words were selected so that they would be familiar to children as young as 4 –years old. The number of words presented was determined based on previous studies showing that 5- to 6-year-olds are able to recall 2-3 items correctly and that 8-year-olds can recall 3-4 items correctly in a WM task (e.g., Barrouillet et al., 2009; Hitch et al., 1989). Since recognition test is less difficult than recall, lists of three words were presented. Children performed the immediate recognition test after a short retention delay. In one condition, they performed a concurrent articulation task during the retention interval to prevent rehearsal, while in the other both rehearsal and refreshing were preserved. Only the use of rehearsal was manipulated in this experiment since this is the most influential mechanism in preventing false memory at short term in adults (Abadie & Camos, 2019).

## **Method**

### **Sample size**

We computed a meta-analytic effect size across studies that reported an interaction between age and the availability of WM maintenance mechanisms on recall performance in children aged 5-14 years using a procedure similar to the present one (Barrouillet et al., 2009; Camos & Barrouillet, 2011). The meta-analytic effect size was  $\eta_p^2 = .131$ ; it was calculated by weighting the  $\eta_p^2$  of the interaction obtained in each study by the number of participants in the study. Power analysis using G\*Power (Faul, et al., 2009) indicated that an overall sample of 56 participants would be needed to achieve an 80% power. In the experiments reported here, we systematically collected larger samples to accommodate for potential data loss. This study was not preregistered.

### **Participants**

Sixty-seven French kindergartners and 3<sup>rd</sup> graders were recruited from several preschools and elementary schools, respectively, in France to participate to the study. Two of

them were excluded from the data analysis because they did not understand the instructions. Out of the 65 remaining children, 32 were kindergartners (the 4-year-old group; 18 girls; *mean age* = 52.92 months; *SD* = 3.16 months) and 33 were 3<sup>rd</sup> graders (the 8-year-old group; 19 girls; *mean age* = 102.3 months; *SD* = 4.10 months). All children were native French speakers. Although we were not able to collect any other demographic variables, we suspect the majority of participants to be from European origin and coming from medium to medium-high socioeconomic backgrounds given the location where the data was collected. The study was conducted in accordance with the local institutional guidelines and APA Ethics Code. For each participant, a written parental consent was obtained. Ethic approval was obtained for both experiments from the institutional review board of Aix-Marseille University (“Working memory and false memories during childhood”, protocol number: 2019-12-12-003).

### **Material**

For the study phase, two subsets of 18 lists were created corresponding to the following thematic categories: circus, farm, school, supermarket, bedroom, soccer game, pirates, fire station, birthday party, Christmas, princesses, beach, zoo, forest, hospital, bathroom, snack time, kitchen (see <https://osf.io/9gfmd/> for the lists development). Each list included three memory items and the average word frequencies of the two subsets did not differ significantly ( $t = .08$ ,  $p = .94$ ,  $d = .02$ ). Children were exposed to one of two subsets.

For the recognition phase, two additional lists of 18 words were constructed. Each of them included six targets (i.e., words from the presented subset), six related distractors (i.e., unrepresented words from the other subset), and six unrelated distractors (i.e., unrepresented concrete words that are not related to any of the thematic categories; the same unrelated distractors were used in both lists). For each of the 18 three-word lists studied, a target probe, a related distractor or an unrelated distractor from the two recognition lists was presented at test. The type of probe presented at test was counterbalanced for each studied list.

Memory items were recorded via ‘audacity’ software (Audacity, 2016) to be auditory presented. Each word was associated with its representation drawn in color. Drawings were copyright-free images found on the internet, pre-selected by the authors of this study to match the words they were to represent. Additionally, three training lists in which all the words pertained to the category “vegetables” were created in the same manner.

### **Procedure**

The study took place in a quiet room of the school. Children were asked to memorize the lists of words related to each category presented as a ‘world’ (e.g., ‘the world of pirates’; ‘the world of princesses’) for a subsequent recognition test. In each trial, children first heard the name of the category, then the 3 items of a list were sequentially presented on screen for 2250 ms per word with a 250 ms interstimulus interval (ISI)<sup>2</sup>. Words were presented in one of three predefined orders. The three presentation orders were counterbalanced across participants.

Next, for half of the 18 trials (i.e., with articulatory suppression), children had to repeat the syllable “ba” aloud at a pace of 1 syllable per second throughout a 6 s retention interval. The experimenter repeated the syllables simultaneously to encourage the children to maintain the repetition until the end of the retention interval. For the other half of the trials (i.e., without articulatory suppression), children remained silent. In each condition (i.e., with and without articulatory suppression), the lists were presented in random order, the conditions as a block, and their order was counterbalanced across participants.

Then, at the end of the trial, children were presented with a probe (i.e., a target, a related or an unrelated distractor) and asked to judge whether or not the probe was in the studied list (Figure 1). A “yes” response was classified as a “target” response. If a “no”

---

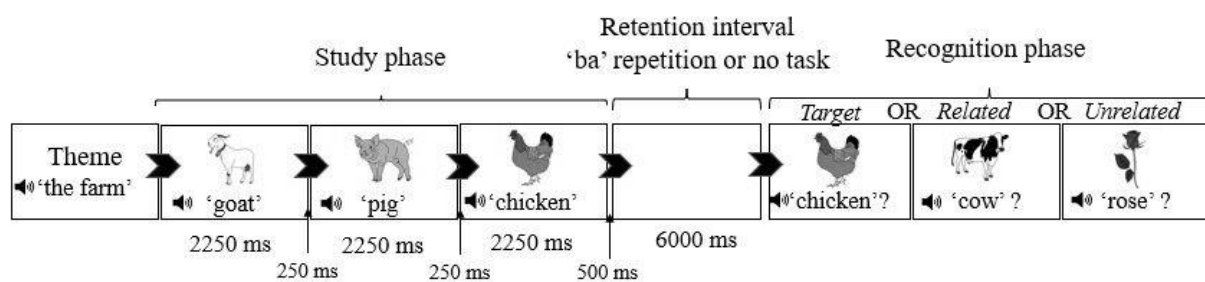
<sup>2</sup> Presentation times of memory items and duration of the retention interval were mapped on previous studies examining WM in children (Barrouillet et al., 2009; Camos & Barrouillet, 2011; Fitamen et al., 2019; Oftinger & Camos, 2018)

response was given, the experimenter would then ask whether or not this probe could belong to the studied category. A “yes” response to this second question was classified as a “related” response, and a “no” as an “unrelated” response. This procedure inspired by the SCR paradigm (Stahl & Klauer, 2008) was used to estimate the contribution of gist and verbatim representations to recognition performance.

A training phase preceding the experiment, in which the two conditions (with and without concurrent articulation) were seen in a random order, was presented to the children. At the end of the experiment, children were rewarded with a diploma.

**Figure 1.**

*Illustration of one trial of the Brown-Peterson task in Experiment 1*



*Note.* Images were retrieved from “Multipic” open access database (Duñabeitia et al., 2018).

**Results**

Data for all experiments are available at <https://osf.io/9gfmd/>

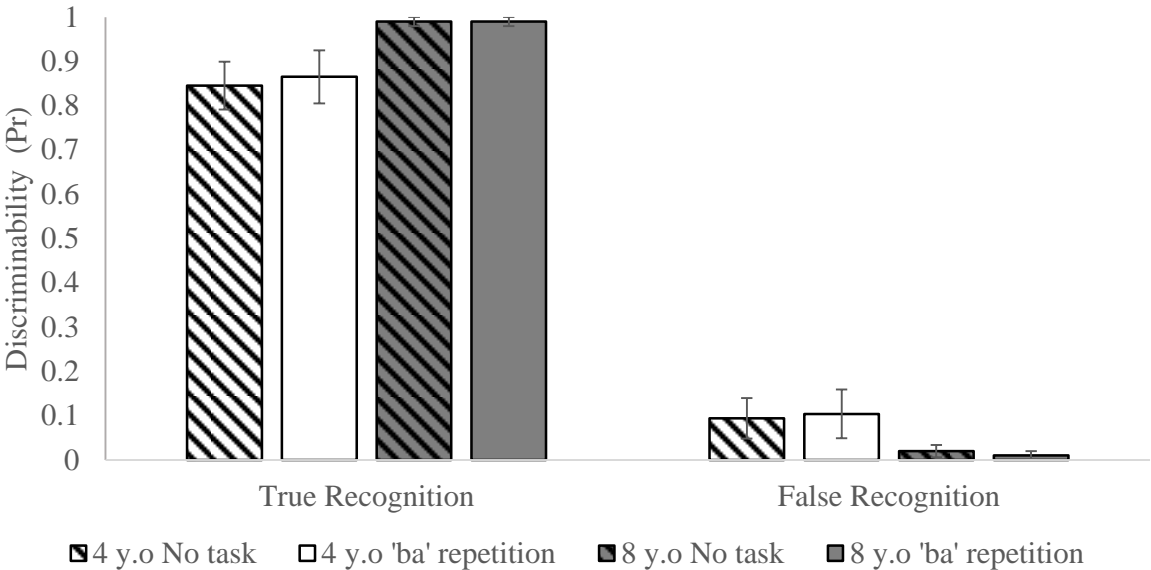
**Memory accuracy**

We computed discriminability indexes ( $Pr$ ) to eliminate potential developmental differences in ‘yes-saying’ bias (i.e., the tendency of young children to answer ‘yes’ to a question more often than older children, e.g., Otgaar et al., 2014). True and false recognition were conditionalized by subtracting the baseline false recognition rate of unrelated distractors (‘yes’ responses to unrelated distractors) from the rate of correct recognition of targets (‘yes’ responses to targets; true recognition) and from the rate of false recognition of related

distractors as targets (“yes” responses to related distractors; false recognition). True and false recognition data are shown in Figure 2.

Mixed-measure ANOVAs were conducted on true and false recognition with age group as between-subject factor (4 vs. 8 years) and type of concurrent task performed during the retention interval (“ba” repetition vs. no task) as within-subject factor. As expected, the main effect of age on true recognition was significant ( $F(1, 63) = 5.81, p = .019, \eta_p^2 = .084, 90\% \text{ CI} [.008; .204]$ ). Eight-year-olds made more true recognition than 4-year-olds. By contrast, age had no significant effect on false recognition ( $F(1, 63) = 3.48, p = .067, \eta_p^2 = .052, 90\% \text{ CI} [0; .161]$ ). Contrary to our prediction, the false recognition rate tended to be even higher in 4- than in 8-year-olds. However, a Bayesian paired-sample t-test provided weak evidence for an age-related difference ( $\text{BF}_{10} = 1.09$ ). The presence of a concurrent articulation during the retention interval did not significantly impact correct nor false recognition ( $F_s < .09$ ). No interaction with age group was significant ( $F_s \leq .18$ ).

**Figure 2.**  
*True and false recognition accuracy as a function of age group and type of concurrent task in Experiment 1.*



*Note.* Error bars represent standard errors.



## **Gist and verbatim memory**

We used the SCR (Stahl & Klauer, 2008) of the FTT to compute four memory parameter estimates and two guessing parameters from the classification of responses as “target”, “related” or “unrelated” for each probe type. Two parameters  $V$  represent the probability of retrieving a verbatim trace of a target when either a target probe ( $V_t$ ) or a related distractor ( $V_r$ ) was presented at test. Verbatim memory traces lead to correct identification of target and related probes. The two parameters  $G$  represent the probability of participants retrieving a gist trace of a target for either a target probe ( $G_t$ ) or a related distractor ( $G_r$ ), given that they have not retrieved its verbatim trace. When participants retrieve gist memory, they cannot remember whether the probe itself or a related word with the same gist had been presented at study. This leads them to guess whether the probe is a target (with probability  $a$ ) or a related distractor (with probability  $1 - a$ ). Therefore, gist retrieval can lead to correct identification of target and related probes or to error in identification of a target as a related distractor or of a related probe as a target item. When neither verbatim nor gist is available, a participant can still guess that the probe meaning is old with the probability  $b$ . The decision between “target” and “related” responses is again modelled by the parameter  $a$ . Otherwise, the participant guesses that the probe is new with the probability  $1 - b$ . Verbatim and gist traces do not intervene when responding to unrelated distractors because these probes do not trigger the retrieval of verbatim or gist representations in the study phase. Therefore, the responses to unrelated distractors are based entirely on  $b$ . The processing tree representation of the model is available in Abadie and Camos' (2019, Figure 3) article and the model equations are provided in the appendix of Stahl and Klauer's (2008) article.

Verbatim, gist and bias parameter estimates are given in Table 1. We used MultiTree Software (Moshagen, 2010) to perform parameter estimations and hypotheses tests. To test the goodness of fit of our data with the model, we tested whether there were significant

violations of the model's four inequality constraints, following Stahl and Klauer's explanations (2008) (see model fit section). Sixteen tests were conducted, none of them revealed a significant violation of the model.

**Table 1.**

*Estimates for the parameters of the simplified conjoint recognition model as a function of age group and type of concurrent task in Experiment 1*

Parameter	4 y.o		8 y.o	
	No Task	'Ba' repetition	No Task	'Ba' repetition
$V_t$	.86 [.39, 1.33]	.88 [.78, .97]	.99 [.97, 1.01]	.99 [.97, 1.01]
$V_r$	.47 [-.94, 1.88]	.52 [.18, .86]	.00 [-3.00, 3.00]	.00 [-4.52, 4.52]
$G_t$	1.00 [-3.08, 5.08]	.80 [.53, 1.07]	1.00 [-1.21, 3.21]	1.00 [-1.18, 3.18]
$G_r$	.79 [.24, 1.34]	.82 [.65, .99]	.97 [.86, 1.07]	.98 [.87, 1.08]
$b$	.16 [.05, .26]	.13 [.07, .20]	.11 [.05, .17]	.10 [.04, .16]
$a$	.33 [-.73, 1.39]	.38 [.12, .65]	.02 [-.04, .08]	.01 [-.03, .05]

Note. These estimates are the proportion of responses to the target, related and unrelated probes in the recognition task based on verbatim memory, gist memory or guessing. 95% confidence intervals are in brackets.

As expected, verbatim memory for targets was greater in 8- than in 4-year-old children ( $\Delta G^2(2) = 20.23, p < .001$ ). Neither verbatim memory for related distractors nor gist memory for targets or for related distractors were significantly impacted by age ( $\Delta G^2(2) = .56, p = .76$ ;  $\Delta G^2(2) = .37, p = .83$ ;  $\Delta G^2(2) = 2.61, p = .27$ , respectively). As mentioned above, according to the SCR model, the retrieval of gist traces occurs when verbatim traces of the studied items are not accessible (i.e.,  $1 - V_t$  for target probes or  $1 - V_r$  for related probes) and can result in two types of responses: the "target" response with probability  $(1 - V_t) \times G_t \times a$  for target probes and  $(1 - V_r) \times G_r \times a$  for related probes, or the "related" response with probability,  $(1 -$

$V_t) \times G_t \times (1 - a)$  for target probes and  $(1 - V_r) \times G_r \times (1 - a)$  for related probes. Thus, considering responses given to related probes on the basis of gist retrieval, it can be noted that gist trace retrieval led 8-year-olds to correctly identifying related distractors in most cases (96%). By contrast, 4-year-olds identified related distractors either as related distractors (26% of cases) or as targets (14% of cases) on the basis of gist memory. Interestingly, 4 year-olds did not guess significantly more often than 8-year-olds that an item meaning was old ( $b$ ;  $\Delta G^2(2) = 1.42, p = .49$ ). However, when they identified an item as old on the basis of guessing or gist memory, 4-year-olds were more likely to identify it as a target than as a related distractor ( $a$ ;  $\Delta G^2(2) = 12.49, p = .002$ ). The concurrent task manipulation had no significant impact on the memory and guessing parameters ( $\Delta G^2s \leq 3.01$ ).

## **Discussion**

As expected, true recognition rate increased with age. Contrary to our predictions, false recognition did not vary significantly as a function of age. Interestingly, in our short-term DRM-like paradigm, 4-year-olds tended to make as many (or even more) false recognitions of related distractors as 8-year-olds. Moreover, memory accuracy of children in both age groups was not significantly affected by articulatory suppression introduced during the retention interval. The analysis of verbatim and gist representations provided a very good account of the recognition performance.

First, verbatim memory increased with age, which perfectly fits the increase in true recognition. Second, gist memory was not significantly affected by age, which directly explains the absence of significant age-related differences in false recognition rates. There was, however, a qualitative difference in the use of gist traces between 4- and 8-year-olds. Indeed, gist retrieval primarily led 8-year-olds to correctly identify related probes as such, whereas it led 4-year-olds to produce both correct and false recognitions. In addition, the tendency of younger children to judge probes with familiar meaning as targets (parameter  $a$ )

seems to be explained by a different use of gist memory rather than by response bias (i.e., the probability of guessing that an item is either a target or a related probe, which is modeled by parameter  $b$ ). Indeed, the parameter  $b$  did not differ between age groups and only accounted for 5% of the target responses in younger children. Hence, it seems that even 4-year-old children were able to process the gist of our short DRM-like lists. Although it can be surprising that such young children were able to extract the gist of DRM lists, it should be remembered that our lists were specifically tailored for them, which could have facilitated gist encoding. Nevertheless, 4-year-olds also appeared to use gist traces differently than 8-year-olds.

In the present experiment, neither memory accuracy nor verbatim and gist representations in 8-year-olds were affected by our manipulation of rehearsal. This might suggest that older children have used strategies other than rehearsal. Moreover, the task was easier for older than younger children as list lengths were identical between age groups. This could explain the absence of a significant difference in false recognition rate between the age groups despite a qualitatively different use of gist memory. Hence, in Experiment 2, our manipulations of the concurrent task targeted not just rehearsal, but also refreshing, while adjusting the difficulty of the task to age group.

## **Experiment 2**

The method of Experiment 2 was similar to that of Experiment 1, with the exception of the number of memory words and their presentation time being adapted to the age of the children, as older children have larger WM capacity and faster processing speed (Camos & Barrouillet, 2018; Cowan, 2014, for reviews). Moreover, the memory items were presented only auditorily to facilitate the use of rehearsal (e.g., Baddeley, 1986; Henry, 2012). As in Experiment 1, we contrasted two conditions that differed in the concurrent task implemented during retention interval. However, in this study the availability of the two WM maintenance

mechanisms, rehearsal and refreshing, was manipulated. In one condition, the use of both WM mechanisms was impaired by a concurrent task involving both articulatory suppression and high attentional demand. Children were asked to name aloud the color of smileys while pressing a corresponding-colored key. This concurrent task is highly attention-demanding for children and prevents the use of rehearsal (Barrouillet et al., 2009; Camos & Barrouillet, 2011; Fitamen et al., 2019). In the other condition, the concurrent task was a simple detection task performed silently. This task did not involve any concurrent articulation and had a low attentional demand, thus allowing the use of the two WM maintenance mechanisms. Finally, a delayed recognition test was added to assess true and false recognition at longer delays in order to compare our results with those obtained in the classical DRM paradigm.

## **Method**

### **Participants**

Thirty-three French kindergartners (4-year-old group; 19 girls; *mean age* = 57.96 months, *SD* = 3.06 months) and 34 3<sup>rd</sup> graders (8-year-old group; 19 girls; *mean age* = 102.26 months, *SD* = 4.10 months) were recruited from several schools in France. All children spoke French as their first language and did not participate to Experiment 1. As in the first experiment, the majority of participants were assumed to be mainly from European origin and coming from medium to medium-high socioeconomic backgrounds given the location where the data was collected. For recruitment purposes, parents were sent a written description of the study and signed a consent form. Data of one kindergartner were excluded from the analysis because his performance in the high-demanding concurrent task was below two standard deviations of the group's average performance. Moreover, due to technical problems, two 3<sup>rd</sup> graders did not achieve the delayed recognition test, and their data were therefore not included in the analysis of this test.

### **Material**

Word lists in this experiment were similar to the ones used in Experiment 1. The two subsets from Experiment 1 were mixed to get 6 words per thematic category. These lists were validated by two tasks conducted prior to the present experiment (see the OSF page).

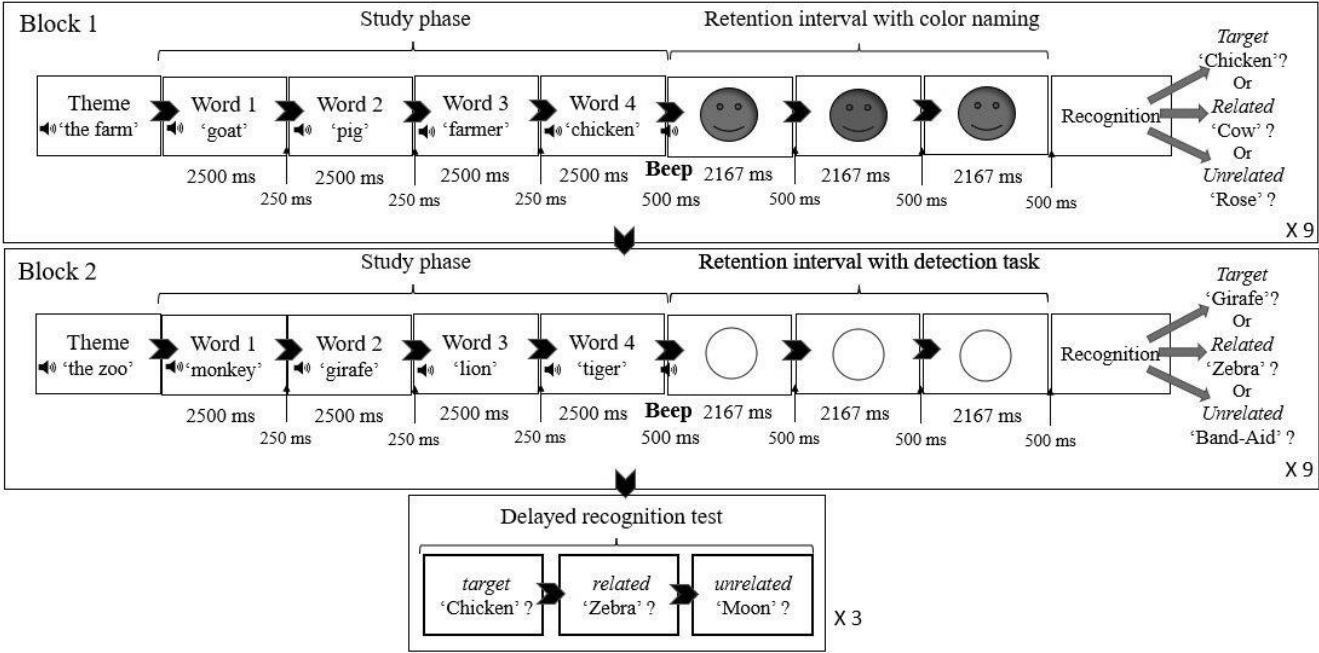
### **Procedure**

The procedure was similar to that of Experiment 1, but changes were made to adjust the task difficulty to each age group (Figure 3). In each trial, children first heard the name of the category, then four or five words were presented sequentially and in a random order on screen to 4- and 8-year-old children, respectively. Words were presented for 2000 ms for 8-year-olds and 2500 ms for 4-year-olds with an ISI of 250 ms in both age groups. Next, a 500 ms “beep” sound signaled the start of the concurrent task. The attentional demand of the concurrent task was adjusted to each age group based on the results of the titration procedure described on the OSF. In order to vary the availability of refreshing, the concurrent task was either high or low attention demanding. In the high attention demanding condition, three or five smileys appeared sequentially on screen for 2167 ms or 1100 ms followed by an ISI of 500 ms for the 4- and 8-year-old children, respectively. To prevent the use of rehearsal, children had to name aloud the color of each smiley while pressing a corresponding-colored key. In the low attention demanding condition, circles appeared sequentially on the screen and children had to press the space bar for each one of them. The number of circles was equal to the number of smileys presented in the high demanding concurrent task. Conditions were presented in blocks, the order of which was counterbalanced across participants. At the end of all trials, children completed a delayed recognition task. Six targets, 6 related distractors and 6 unrelated distractors were sequentially presented in a random order. Children had to identify each of them using the same procedure as in the immediate recognition task.

A training phase preceded the experimental phase. First, children completed the two concurrent tasks alone on 12 stimuli each. Then, they performed two trials of the Brown-Peterson task for each condition.

**Figure 3.**

*Illustration of trials in Experiment 2.*



*Note.* Eight-year-old group was shown one extra-word during the study phase and had to process one more stimulus than the 4-year-olds during the retention interval. The order of presentation of the blocks was counterbalanced among participants.

**Results**

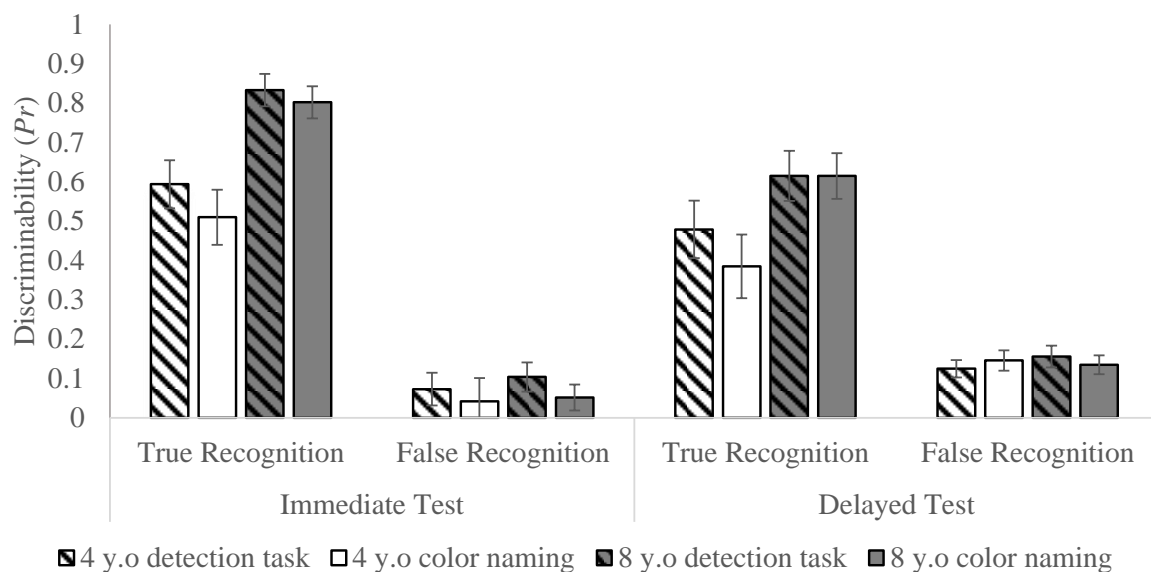
Similar analyses as in Experiment 1 were conducted on both immediate and delayed recognition tests. Mean performance in both concurrent tasks were above 80% accuracy in each age group (detection  $M-4years = 85.5\%$ ,  $SD = 9.3$ ;  $M-8years = 93.2\%$ ,  $SD = 7.8$ ; color naming:  $M-4years = 82.9\%$ ,  $SD = 13.06$ ;  $M-8years = 87.7\%$ ,  $SD = 10.0$ ), which showed that children followed the instructions.

**Memory accuracy**

Mixed-measure ANOVAs were conducted on  $Pr$  indexes for true and false recognition with age group as between-subject factor (4 vs. 8 years old) and type of concurrent task (detection vs. color naming) as within-subject factor for each probe type and time of test separately. There was a significant main effect of age on true recognition rate in both immediate and delayed tests ( $F(1, 64) = 18.31, p < .001, \eta_p^2 = .23, 90\% \text{ CI } [.087; .354]$ ;  $F(1, 64) = 4.85, p = .03, \eta_p^2 = .07, 90\% \text{ CI } [.003; .185]$ , respectively). As predicted, the true recognition rate was higher in 8-year-olds than 4-year-olds in both tests (Figure 4). However, as in Experiment 1, false recognition rate was not significantly affected by age in either the immediate or delayed tests ( $F(1, 64) = .16, p = .69, \eta_p^2 = .003, 90\% \text{ CI } [0; .055]$ ;  $F(1, 64) = .05, p = .83, \eta_p^2 = 7.58 \times 10^{-5}, 90\% \text{ CI } [0; .037]$ , respectively). Moreover, a Bayesian t-test provided substantial evidence for an absence of age-related difference in false recognition in both tests ( $\text{BF}_{10} = .27$ ;  $\text{BF}_{10} = .26$ , respectively). Finally, the type of concurrent task did not significantly affect correct nor false recognition and there was no interaction with age ( $F_s \leq 1.52$ ).

**Figure 4.**

*True and false recognition accuracy as a function of age group and type of concurrent task for immediate and delayed tests in Experiment 2.*





*Note.* Error bars represent standard errors.

### **Gist and verbatim memory**

Verbatim, gist and bias parameter estimates for immediate and delayed recognition tests are given in Table 2. As in Experiment 1, we tested whether there were significant violations of the model's four inequality constraints. Sixteen tests were conducted for the immediate and delayed test, none of them revealed a significant violation of the model.

**Table 2.**

*Estimates for the parameters of the SCR model as a function of age groups and concurrent tasks in both immediate and delayed tests in Experiment 2*

Time of test	Parameter	4 y.o		8 y.o	
		Detection task	Color Naming	Detection task	Color Naming
Immediate	$V_t$	.67 [.51, .83]	.60 [.41, .78]	.79 [.70, .88]	.80 [.71, .88]
	$V_r$	.25 [-.02, .52]	.22 [-.02, .46]	.00 [-1.25, 1.25]	.00 [-2.64, 2.64]
	$G_t$	.77 [.54, 1.00]	.68 [.41, .95]	.95 [.84, 1.05]	.95 [.85, 1.05]
	$G_r$	.65 [.43, .86]	.51 [.25, .77]	.95 [.88, 1.02]	.97 [.88, 1.06]
	$b$	.45 [.35, .55]	.51 [.41, .61]	.11 [.05, .17]	.05 [.01, .09]
	$a$	.46 [.31, .61]	.51 [.37, .65]	.09 [-.03, .22]	.05 [-.08, .18]
Delayed	$V_t$	.54 [.32, .77]	.41 [.12, .71]	.55 [.42, .67]	.57 [.45, .69]
	$V_r$	.00 [-.16, .16]	.01 [-.15, .16]	.00 [-.58, .58]	.00 [-.64, .64]
	$G_t$	.48 [.13, .83]	.39 [.02, .76]	.70 [.52, .87]	.68 [.50, .86]
	$G_r$	.22 [-.04, .48]	.35 [.13, .58]	.74 [.56, .92]	.74 [.55, .93]
	$b$	.47 [.37, .57]	.42 [.32, .52]	.22 [.13, .30]	.22 [.13, .29]
	$a$	.65 [.53, .78]	.72 [.59, .86]	.16 [.04, .27]	.16 [.04, .28]

*Note.* 95% confidence intervals are in brackets.

*Immediate test.* As expected and observed in Experiment 1, verbatim memory for targets was greater in 8- than in 4-year-olds ( $\Delta G^2(2) = 6.72, p = .034$ ). There was a significant effect of age on verbatim memory for targets when the concurrent task was the color naming task ( $\Delta G^2(1) = 4.67, p = .031$ ), but not for the detection task ( $\Delta G^2(1) = 2.05, p = .15$ ).

Moreover, 8-year-olds exhibited greater gist memory for targets than 4-year-olds ( $\Delta G^2(2) = 6.20, p = .04$ ). This effect was also significant when the concurrent task was the color naming task ( $\Delta G^2(1) = 4.08, p = .04$ ), but not for the detection-task ( $\Delta G^2(1) = 2.13, p = .14$ ).

Furthermore, when decomposing the response type (target or related) resulting from the retrieval of gist memory for targets, we observed that 8-year-olds identified targets as related distractors (18%) more often than as targets (1%). By contrast, 4-year-olds equally identified targets as either targets (13%) or related distractors (13%). In addition, older children had greater values of gist memory for related distractors than younger children ( $\Delta G^2(2) = 14.85, p < .001$ ). The effect was present whether the color naming task ( $\Delta G^2(1) = 9.24, p < .002$ ) or the detection task ( $\Delta G^2(1) = 8.43, p < .004$ ) was used as the concurrent task. As in Experiment 1, gist retrieval led 8-year-olds to more often correctly identify related distractors as such (89%) rather than as targets (7%). By contrast, 4-year-olds identified related distractors as related (23%) as often as targets (21%) based on gist traces retrieval. Finally, the two guessing parameters ( $a$  and  $b$ ) were significantly greater in younger children ( $\Delta G^2(2) = 15.61, p < .001$ ;  $\Delta G^2(2) = 86.47, p < .001$ , respectively). Four-year-olds correctly identified targets on the basis of guessing (parameter  $b$ ) in 2% of cases and they identified related distractors as such in 8% of cases and as targets also in 8% of cases. This never occurred in 8-year-olds for whom this parameter virtually did not contribute to these types of responses. Moreover, 4-year-olds were more likely than 8-year-olds to guess that an unrelated distractor was a target or a related probe (24% and <1% respectively).

*Delayed test.* Verbatim and gist memory for targets were lower in this test compared to the immediate test and they were not significantly affected by age ( $\Delta G^2(2) = .68, p = .71$ ;  $\Delta G^2(2) = 3.63, p = .16$ , respectively). Gist memory for related distractors was significantly higher in 8- than 4-year-olds ( $\Delta G^2(2) = 22.30, p < .001$ ). Gist retrieval led 8-year-olds to more often correctly identifying related probes as such (62%) rather than as targets (12%). By contrast 4-year-olds were more likely to identify falsely related probes as targets (19%) than as related distractors (9%) based on gist retrieval. As in the immediate test, the two guessing parameters ( $a$  and  $b$ ) were greater in the younger children group ( $\Delta G^2(2) = 59.84, p < .001$ ;  $\Delta G^2(2) = 22.74, p < .001$ , respectively). Parameter  $b$  accounted for 9% of correct identification of targets, 10% of correct identification of related distractors and 22% of false identification of related distractors as targets in 4-year-olds. Parameter  $b$  contributed only marginally to the responses to target and related probes of 8-year-olds (at most 4% for the correct identification of related distractors). In addition, 4-year-olds were more likely than 8-year-olds to guess that an unrelated distractor was either a target or a related probe (22% and 11% of cases, respectively). All other effects were not significant.

## **Discussion**

Results of Experiment 1 were replicated even though the task difficulty (i.e., memory load and concurrent task load) was adapted for each age group. In both the immediate and the delayed tests, the true recognition rate was higher in 8- than 4-year-olds, whereas the false recognition rate did not differ as a function of age group. In both age groups, memory accuracy was not affected by the manipulation of WM mechanisms availability. As in Experiment 1, the analysis of the contribution of verbatim and gist representations provided a good account of the recognition performance.

In the immediate test, verbatim and gist memory for targets were greater in 8- than 4-year-olds, particularly when the concurrent task was more demanding. This result accounts

for the increase in true recognition with age. Older children also exhibited greater gist memory for related distractors than 4-year-olds. Here, gist memory retrieval led 8-year-olds to more often correctly judge that the item they were judging was a related distractor. By contrast, gist memory in 4-year-olds led them to make more false identification of related distractors as targets. These qualitative differences in the use of gist memories explain why, although they retrieved more gist memories, older children did not make more false memories than younger children. The same pattern of results for gist memory for related distractors was found in the delayed test. Congruently with Experiment 1, these results suggest that gist-based false memories can occur in children as young as 4 years old.

Finally, as in Experiment 1 and despite the adaptation of the task difficulty to the children's age group, recognition performance was not affected by the type of concurrent task. This could be due to the auditory presentation of words, which may have encouraged the use of rehearsal and to the fact that uttering the names of colors did not sufficiently block the use of this mechanism. We conducted an additional experiment, the method and results of which are described in detail on the OSF, in which we intended to minimize the use of rehearsal and encourage the use of refreshing to maintain the information in WM. However, false recognition rate was at near floor in this additional experiment, rendering hazardous any conclusion. This could have resulted from the use of pictures; previous studies showed that visual supports weaken the DRM illusion in older children (e.g., Brainerd et al., 2018; see Brainerd & Reyna, 2012, for a review). These different points were addressed in the general discussion.

### **General Discussion**

The present study examined how the age-related changes in verbatim and gist representations, and in the use of WM maintenance mechanisms affect children's correct and false recognition in a WM task. The availability of WM maintenance mechanisms did not

affect recognition performance in any of the experiments, contrary to our expectations. However, we demonstrated two robust phenomena. First, 4- and 8-year-old children made false recognition of related distractors in both immediate and delayed recognition tests. Second, congruently with the FTT's predictions, false recognitions were underpinned by gist memory in both age groups. We will discuss in turn the false recognition effect obtained in 4- and 8-year-olds, the absence of significant effect of the manipulation of WM maintenance mechanisms in young children's recognition, and the role of gist memory in the formation of children's short- and long-term false memories.

### **False recognition effect in 4- and 8-years old children**

Our study demonstrated that semantic distortions could occur in a WM task in 4- and 8-years old children. Previous studies have shown that semantically related distractors can be falsely recognized by young adults within seconds of the study phase (e.g., Abadie & Camos, 2019; Atkins & Reuter-Lorenz, 2008). To our knowledge, our study is the first to examine and evidence this phenomenon in children.

In addition, although false memories occur at short term in children, their incidence did not vary with age as in the classical DRM illusion. Younger children are as prone to false memory at short term as their older counterparts. These findings are both interesting and unexpected since developmental LTM studies using the DRM paradigm typically show an age-related increase in the occurrence of false memories, with younger children being less susceptible to DRM false memories than older children (e.g., Anastasi & Rhodes, 2008; Brainerd et al., 2002c). At first, this might be taken as evidence that it is much easier for young children to connect and use semantic gist in WM tasks than in LTM tasks.

However, it is also possible that giving the theme of each list prior to the word presentation might have facilitated the retrieval of the gist of the lists. Studies indeed showed that gist-cuing before the presentation of DRM lists increased false recall and recognition of

critical lures in children as young as 5 years old (see Brainerd et al., 2008, for a review). In our study, this may have particularly benefited young children's ability to process the meaning of the items, favoring the creation of visual scenes based on the items. For example, the presentation of 'the farm' theme may prompt the image of a farm in which each element (pig, farmer, etc.) is easily integrated, as this type of visual scene is very common in 4-year-olds' books and cartoons and probably more so than for 8-year-olds. Furthermore, the presentation of a theme word may have a different impact on the processing of the common meaning (i.e., gist) from multiple items in young and older children. Indeed, older children are better at extracting the gist from multiple items (e.g., Brainerd et al., 2002c) and the occurrence of false memory in these children may not be influenced by the presentation of the theme word. By contrast, the theme word likely scaffolded the retrieval of common meaning in younger children, reducing the age-related difference in the incidence of false memories. This interpretation is partially supported by previous studies. For instance, Brainerd et al. (2004) have shown that telling 5- and 11- year-olds that the studied list exemplars would all be part of a specific taxonomic category before list-presentation (e.g., "all the words on the next list will be animals"; gist-cuing condition) increased false recognition in both age-groups. In the same vein, Holliday et al. (2008) showed that the increase in false recall between 7- and 15-year-olds was smaller in a gist-cuing condition compared to a control condition. Hence, these findings suggest that age-related increases in false memory may be attenuated by gist cuing in our study, but gist cuing might not be the only reason.

Other studies have indeed shown that the standard developmental trend in LTM tasks can be attenuated or even reversed when using modified versions of the DRM paradigm in which the DRM words are embedded in stories (e.g., Ghetti, et al., 2002). Story contexts enhanced 5-year-olds' vulnerability to the DRM illusion (e.g., Dewhurst et al., 2007; Howe & Wilkinson, 2011). Using scene-like visual stimuli that are associatively related with each

other also resulted in higher false memory rates in 7- and 8-year-old children than in adults (Otgaar et al., 2014). The word lists used in the present study were specifically created for young children, i.e., they contained themes that were well known and relevant to them. Carneiro et al. (2007) showed that when the studied lists were adapted to age groups by selecting the words rated as most associated to the critical lure according to the target group, false memories increased in the youngest group (3–4-years old) and the usual developmental increase trend was reduced. Hence, list of words designed for and by children could also be responsible for such a high rate of false memories in the youngest group. However, other studies have documented an increase in false memories with age using DRM lists created by children (e.g., Anastasi & Rhodes, 2008, Exp. 3; Metzger et al. 2008). It is worth noting that the latter studies used classical DRM themes, which was not the case in our study in which we used themes that are well known and more related to young children's daily life (such as the princesses, the farm, etc.) than DRM themes. In addition, the use of shorter list lengths in our WM task may have contributed to differences with the above-mentioned studies which implemented the long list lengths, typical of research of false memory in LTM. Further studies should aim at disentangling the impact of theme word presentation and of the type of lists used for generating false memories in WM tasks across different age groups.

### **The impact of WM maintenance on recognition performance**

Previous findings have suggested that children from age 7 onwards can use rehearsal and/or refreshing to maintain information in WM, whereas younger children do not spontaneously use these mechanisms (e.g., Camos & Barrouillet, 2011; Henry, 2012; Oftringer & Camos, 2018; Tam et al., 2010). In the present study, as expected, 4-year-olds did not seem to engage in information maintenance strategies. Indeed, their responses to recognition tests were not impacted by the type of concurrent task in any of the experiments. Similarly, the 8-

year-olds' responses were also not impacted by the type of concurrent task, which was rather unexpected.

There are two possible reasons for this. First, it is possible that our concurrent tasks did not properly suppress rehearsal of memory words or were not attentionally demanding enough to prevent refreshing. However, the concurrent tasks used in our experiments were similar to those implemented in previous studies in which they had affected the memory performance of children of the same age (Barrouillet et al., 2009; Camos & Barrouillet, 2011; Oftinger & Camos, 2018). Nevertheless, it may be noted that Langerock et al. (2020) showed that, in adults, the manipulation of concurrent attentional demand (i.e., the cognitive load effect) has a weaker effect (or even none) in Brown-Peterson tasks than in complex span tasks used in the previously mentioned studies. Thus, one might expect that a stronger manipulation of cognitive load would be required to observe its effect in a Brown-Peterson task, especially in children. However, the princeps studies showing short-term false memories in adults implemented a Brown-Peterson task (Abadie & Camos, 2019; Atkins & Reuter-Lorenz, 2008), and therefore called for the use of this type of task in the first study in children, for the sake of comparisons.

Second, another explanation could be that 8-year-olds did not use WM maintenance mechanisms in our recognition task. As recently shown by Uittenhove et al. (2019) in adults, recognition test responses are supported by passive storage of memory items in LTM. This phenomenon is likely to have been even stronger in the present study among older children for whom the recognition test may have been perceived as particularly easy and effortless. Moreover, studies showed that compared to adults, children tend to rely more on reactive control, which involves passively engaging in a task (Chevalier, et al., 2014; Lucenet & Blaye, 2014; Munakata, et al., 2012). In contrast, proactive control is required to actively maintain information in WM and implement WM maintenance mechanisms (Braver 2012). It



should be noted that previous experiments showing effects of WM maintenance mechanisms suppression in children's memory performance have used recall tasks and not recognition tasks (e.g., Barrouillet et al., 2009; Camos & Barrouillet, 2011; Oftinger & Camos, 2016; Tam et al., 2010). However, in our study, recognition was preferred over recall, because the use of a recognition task seemed more appropriate in 4-year-old children to avoid floor effect. Finally, to assess gist and verbatim, we implemented the SCR model, which can only apply to recognition tasks.

### **The role of gist memory in false memories in children**

First of all, it is important to point out that, in all the experiments, the SCR model of the FTT fitted the data of both age groups very well, demonstrating that children in both age groups understood and responded well to our recognition tasks.

The FTT predicts that gist memory improves during childhood, because the ability to process the meaning of items and to understand the relationships between different items improves (e.g., Brainerd & Reyna, 2015). Improvements in gist memory increase the tendency to falsely accept semantically related lures as having been previously studied in LTM (e.g., Brainerd et al., 2002c). In the present study, both 4- and 8-year-olds seem to be able to make meaning connections among the items of our short DRM-like lists, which led them to falsely recognize related distractors in the immediate and delayed tests. Although there were no significant differences in gist memory use between the two age groups in Experiment 1, we found an increase in gist memory with age in Experiment 2. In the latter experiment, memory lists were presented orally, which is the most common methodology used to study the DRM illusion. Studies showed that using visual supports, as it was the case in Experiment 1, weaken the illusion in older children (e.g., Brainerd et al., 2018; see Brainerd & Reyna, 2012, for a review). As shown in adults (e.g., Abadie et al., 2017), the

format in which items are presented seems likely to modulate the tendency to form and retrieve gist traces in children.

However, these methodological differences cannot explain why gist memory increases with age, while older children do not make more false memories than younger children. A better explanation might lie in the qualitatively different use of gist memory depending on age. Combining the results of both experiments, we found that 8-year-olds used gist memory in most cases to correctly identify related distractors as such (i.e., an average of 96% of the time in immediate tests and of 84% of the time in the delayed test). By contrast, gist retrieval led 4-year-olds to identify related distractors either as such (on average, 58% of cases) or as target probes in an undifferentiated manner in the immediate tests and mostly as targets in the delayed test (on average, 70% of cases). Similarly, gist memory for target probes led the 8-year-olds to identify them correctly in the majority of cases (93% and 86% of cases in immediate and delayed tests, respectively), while it led the 4-year-olds to identify them almost as often as target probes (67% and 59%, respectively) as related distractors.

These findings suggest that, although young children were able to make meaningful connections between items of each list, which may explain the absence of age-related difference in false memory rates, there are still qualitative differences between 4- and 8-year-old children in the use of gist memory.

### **Conclusion**

The present study aimed at examining how the development of LTM and WM can impact the emergence of false memory in 4- and 8-year-old children in WM tasks. Despite some evidence in adults (Abadie & Camos, 2019; Atkins & Reuter-Lorenz, 2008), this issue had never been examined in children. This study, in which short DRM-like lists specifically tailored for young children were presented in a Brown-Peterson task, found that younger children are as prone to false memories as their older counterparts in the immediate

recognition test, and in a delayed recognition test. In addition, using the state-of-the-art mathematical model of the FTT, we demonstrated for the first time in the literature that the retrieval of gist traces of the list themes is responsible for the emergence of short-term false memories in 4- and 8-year-olds. Gist memory also underpinned the occurrence of false recognition in the delayed test, replicating the classical results obtained in LTM with the DRM paradigm. These findings suggest that young children are as likely to make gist-based false memories as older children in WM tasks.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## References

- Abadie, M., & Camos, V. (2019). False memory at short and long term. *Journal of Experimental Psychology: General*, *148*(8), 1312–1334.  
<https://doi.org/10.1037/xge0000526>
- Abadie, M., Waroquier, L. (2020). On the memory processes underlying conscious deliberation in complex decision making: the role of verbatim and gist memory. *Psychological Research* *84*, 1714–1722. <https://doi.org/10.1007/s00426-019-01180-8>
- Abadie, M., Waroquier, L., & Terrier, P. (2013). Gist memory in the unconscious thought effect. *Psychological Science*, *24*(7), 1253-1259.  
<https://doi.org/10.1177/0956797612470958>
- Abadie, M., Waroquier, L., & Terrier, P. (2017). The role of gist and verbatim memory in complex decision making: Explaining the unconscious-thought effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(5), 694-705.  
<https://doi.org/10.1037/xlm0000336>
- Anastasi, J.S., & Rhodes, M.G. (2008). Examining differences in the levels of false memories in children and adults using child-normed lists. *Developmental Psychology*, *44*, 889–894.  
<https://doi.org/10.1037/0012-1649.44.3.889>
- Atkins, A.S., & Reuter-Lorenz, P.A. (2008). False working memories? Semantic distortion in a mere 4 seconds. *Memory and Cognition*, *36*(1), 74–81.  
<https://doi.org/10.3758/MC.36.1.74>
- Audacity team (2016). Audacity (version 2.4.2) [Computer software]. Retrieved from <https://www.audacityteam.org/>

- Baddeley, A.D. (1986). *Working memory*. Oxford, UK: Clarendon Press.
- Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction*. Hove, U.K.: Psychology Press.
- Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development: A time-based resource-sharing model account. *Developmental Psychology*, *45*(2), 477-490.  
<https://psycnet.apa.org/doi/10.1037/a0014615>
- Brainerd, C. J., & Reyna, V. F. (2002b). Recollection rejection: How children edit their false memories. *Developmental Psychology*, *38*(1), 156-172. <https://doi.org/10.1037/0012-1649.38.1.156>
- Brainerd, C.J., & Reyna, V. F. (1998). Fuzzy-Trace Theory and Children's False Memories. *Journal of Experimental Child Psychology*, *71*(2), 81–129.  
<https://doi.org/10.1006/jecp.1998.2464>
- Brainerd, C.J., & Reyna, V. F. (2005). *The science of false memory*. New York, NY: Oxford University Press.
- Brainerd, C.J., & Reyna, V. F. (2012). Reliability of children's testimony in the era of developmental reversals. *Developmental Review*, *32*, 224-267.  
<https://doi.org/10.1016/j.dr.2012.06.008>
- Brainerd, C.J., & Reyna, V. F. (2015). Fuzzy-trace theory and lifespan cognitive development. *Developmental Review*, *38*, 89-121.  
<https://doi.org/10.1016/j.dr.2015.07.006>
- Brainerd, C.J., & Reyna, V.F. (2002a). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, *11*(5), 164–169. <https://doi.org/10.1111/1467-8721.00192>

- Brainerd, C.J., Holliday, R.E., & Reyna, V.F. (2004). Behavioral measurement of remembering phenomenologies: So simple a child can do it. *Child Development, 75*(2), 505-522. <https://doi.org/10.1111/j.1467-8624.2004.00689.x>
- Brainerd, C.J., Nakamura, K., Chang, M., & Bialer, D.M. (2019). Verbatim editing: A general model of recollection rejection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(10), 1776–1790. <https://doi.org/10.1037/xlm0000684>
- Brainerd, C.J., Reyna, V.F., & Ceci, S.J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin, 134*, 343–382. <https://doi.org/10.1037/0033-2909.134.3.343>
- Brainerd, C.J., Reyna, V.F., & Forrest, T.J. (2002c). Are Young Children Susceptible to the False–Memory Illusion? *Child Development, 73*(5), 1363–1377. <https://doi.org/10.1111/1467-8624.00477>
- Brainerd, C.J., Reyna, V.F., & Holliday, R.E. (2018). Developmental reversals in false memory: Development is complementary, not compensatory. *Developmental Psychology, 54*(9), 1773-1784. <https://doi.org/10.1037/dev0000554>
- Brainerd, C.J., Reyna, V.F., & Mojardin, A.H. (1999). Conjoint Recognition. *Psychological Review, 106*(1), 160-179.
- Brainerd, C.J., Reyna, V.F., Wright, R., & Mojardin, A.H. (2003). Recollection Rejection: False-memory editing in children and adults. *Psychological Review, 110*(4), 762-784. <https://doi.org/10.1037/0033-295X.110.4.762>
- Braver, T.S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences, 16*, 106–113. <https://doi.org/10.1016/j.tics.2011.12.010>
- Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10*, 12-21. <https://doi.org/10.1080/17470215808416249>

- Buckey v. County of Los Angeles, 968 F.2d 791 (9th Cir. 1992), 90-56333.
- Camos, V. & Barrouillet, P. (2011). Developmental change in working memory strategies: from passive maintenance to active refreshing. *Developmental Psychology*, 47(3), 898-904. <https://doi.org/10.1037/a0023193>
- Camos, V. (2015). Storing verbal information in working memory. *Current directions in Psychological Science*, 24(6), 440-445. <https://doi.org/10.1177/0963721415606630>
- Camos, V. (2017). Domain-specific vs. domain-general maintenance in working memory: reconciliation within the time-based resource sharing model. In. H. Ross (Ed.) *Psychology of Learning and Motivation*, Vol. 67 (pp. 135–171). Cambridge, MA: Academic Press.
- Camos, V., & Barrouillet, P. (2018). Working Memory in Development (1st ed.). Routledge.
- Camos, V., & Portrat, S. (2015). The impact of cognitive load on delayed recall. *Psychonomic Bulletin & Review*, 22, 1029–1034. <https://doi.org/10.3758/s13423-014-0772-5>
- Carneiro, P., Albuquerque, P., Fernandez, A. & Esteves, F. (2007). Analyzing false memories in children with associative lists specific for their age. *Child Development*, 78(4), 1171-1185.
- Ceci, S.J., & Bruck, M. (1998). The ontogeny and durability of true and false memories: A fuzzy trace account. *Journal of Experimental Child Psychology*, 71(2), 165-169. <https://doi.org/10.1006/jecp.1998.2468>
- Coane, J.H., McBride, D.M., Raulerson, B.A., III, & Jordan, J.S. (2007). False memory in a short-term memory task. *Experimental Psychology*, 54, 62–70. <https://doi.org/10.1027/1618-3169.54.1.62>

- Cowan, N. (2014). Short-term and working memory in childhood. In P.J. Bauer & R. Fivush (Eds.), *The Wiley handbook on the development of children's memory* (p. 202–229). Wiley Blackwell.
- Chang, M., & Brainerd, C.J. (2021). Semantic and phonological false memory: A review of theory and data. *Journal of Memory and Language*, 119, 104210. <https://doi.org/10.1016/j.jml.2020.104210>
- Chevalier, N., James, T.D., Wiebe, S.A., Nelson, J.M., & Espy, K.A. (2014). Contribution of reactive and proactive control to children's working memory performance: Insight from item recall durations in response sequence planning. *Developmental Psychology*, 50, 1999–2008. <https://doi.org/10.1037/a0036644>
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58(1), 17–22. <https://doi.org/10.1037/h0046671>
- Dewhurst, S.A., Pursglove, R.C., Lewis, C. (2007). Story contexts increase susceptibility to the DRM illusion in 5-years-old. *Developmental Science*, 10, 374 –378. <https://doi.org/10.1111/j.1467-7687.2007.00592.x>
- Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 71, 808 – 816.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.G. (2009). Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>



- Fitamen, C., Blaye, A., & Camos, V. (2019). The role of goal cueing in kindergarteners' working memory. *Journal of Experimental Child Psychology*, *187*, 104666. <https://doi.org/10.1016/j.jecp.2019.104666>
- Flegal, K.E., & Reuter-Lorenz, P.A. (2014). Get the gist? The effects of processing depth on false recognition in short-term and long-term memory. *Memory & Cognition*, *42*, 701–711. <https://doi.org/10.3758/s13421-013-0391-9>
- Flegal, K.E., Atkins, A.S., & Reuter-Lorenz, P.A. (2010). False memories seconds later: The rapid and compelling onset of illusory recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1331–1338. <https://doi.org/10.1037/a0019903>
- Gathercole, S., Pickering, S., & Ambridge, B. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*, 177–190. <https://doi.org/10.1037/0012-1649.40.2.177>
- Ghetti, S., Qin, J., & Goodman, G.S. (2002). False memories in children and adults: age, distinctiveness, and subjective experience. *Developmental Psychology*, *38*(5), 705–718. <https://doi.org/10.1037//0012-1649.38.5.705>
- Henry, L. A. (2012). *The development of working memory in children*. London: Sage.
- Henry, L.A., Messer, D., Luger-Klein, S., & Crane, L. (2012). Phonological, visual and semantic coding strategies and children's short-term picture memory span. *Quarterly Journal of Experimental Psychology*, *65*, 2033–2053. <https://doi.org/10.1080/17470218.2012.672997>
- Hitch, G.J., Halliday, M.S., Dodd, A., Littler, J.E. (1989). Development of rehearsal in short-term memory: Differences between pictorial and spoken stimuli. *British Journal of Developmental Psychology*, *7*(4), 347–362. <https://doi.org/10.1111/j.2044-835X.1989.tb00811.x>

- Holliday, R.E. Reyna, V.F., & Brainerd, C.J. (2008). Recall of details never experienced: Effects of age, repetition, and semantic cues. *Cognitive Development, 23*, 67-78.
- Holliday, R.E., Brainerd, C.J., & Reyna, V.F. (2011). Developmental reversals in false memory: Now you see them, now you don't! *Developmental Psychology, 47*, 442-449.  
<https://doi.org/10.1037/a0021058>
- Howe, M.L. & Wilkinson, S. (2011). Using story contexts to bias children's true and false memories. *Journal of Experimental Child Psychology, 108*(1), 77-95.  
<https://doi.org/10.1016/j.jecp.2010.06.009>
- Howe, M.L. (2000). *The fate of early memories*. Washington, DC: APA.
- Langerock N., & Vergauwe, E. (2022). A robust cognitive load effect in 8 to 12 year old children? [Poster presentation].  
[https://www.researchgate.net/publication/359210154\\_A\\_robust\\_cognitive\\_load\\_effect\\_in\\_8\\_to\\_12\\_year\\_old\\_children](https://www.researchgate.net/publication/359210154_A_robust_cognitive_load_effect_in_8_to_12_year_old_children)
- Langerock, N., Vergauwe, E., Thom, E., & Oberauer, K. (2020, September 1-3). The effect of cognitive load in working memory tasks: Limited to the complex span paradigm? [Conference session]. 10<sup>th</sup> European Working Memory Symposium, Cardiff, Wales.  
[https://www.youtube.com/watch?v=4gwzJE5qZsw&list=PLbongzUA9knd6D4yfpa4\\_1lg\\_tKNfXQ4Oa&index=4](https://www.youtube.com/watch?v=4gwzJE5qZsw&list=PLbongzUA9knd6D4yfpa4_1lg_tKNfXQ4Oa&index=4)
- Lucenet, J., & Blaye, A. (2014). Age-related changes in the temporal dynamics of executive control: A study in 5-and 6-year-old children. *Frontiers in Psychology, 5*, 1-11.  
<https://doi.org/10.3389/fpsyg.2014.00831>
- Metzger, R. L., Warren, A. R., Shelton, J. T., Price, J., & Reed, A. W. (2008). Do children 'DRM' like adults ? False memory production in children. *Developmental Psychology, 44* (1), 169-181.

- Moshagen, M. (2010). MultiTree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods*, 42(1), 42-54.  
<https://doi.org/10.3758/BRM.42.1.42>
- Munakata, Y., Snyder, H.R., & Chatham, C.H. (2012). Developing cognitive control: Three key transitions. *Current Directions in Psychological Science*, 21, 71–77.  
<https://doi.org/10.1177/0963721412436807>
- Odegard, T.N., Holliday, R.E., Brainerd, C.J., Reyna, V.F. (2008). Attention to global gist processing eliminates age effects in false memories. *Journal of Experimental Child Psychology*, 99, 96-113. <https://doi.org/10.1016/j.jecp.2007.08.007>
- Oftinger, A., & Camos, V. (2016). Maintenance mechanisms in children’s verbal working memory. *Journal of Educational and Developmental Psychology*, 6 (1), 16-28.  
<https://doi.org/10.5539/jedp.v6n1p16>
- Oftinger, A., & Camos, V. (2018). Developmental improvement in strategies to maintain verbal information in working memory. *International Journal of Behavioral Development*, 42(2), 1-10. <https://doi.org/10.1177/0165025416679741>
- Otgaar, H., Howe, M.L., Peters, M., Smeets, T., & Moritz, S. (2014). The production of spontaneous false memories across childhood. *Journal of Experimental Child Psychology*, 121, 28-41. <https://doi.org/10.1016/j.jecp.2013.11.019>
- Peterson, L.R., & Peterson, M.J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58, 193–198.
- Reyna, V.F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology*, 30, 178-191. <https://doi.org/10.1037/0012-1649.30.2.178>

- Reyna, V.F., & Lloyd, F. (1997). Theories of false memory in children and adults. *Learning and Individual Differences, 9*, 95–123. [https://doi.org/10.1016/S1041-6080\(97\)90002-9](https://doi.org/10.1016/S1041-6080(97)90002-9)
- Roediger, H.L., & McDermott, K.B. (1995). Creating False Memories: Remembering Words Not Presented in Lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803–814.
- Rousselle, M., Abadie, M., Blaye, A., & Camos, V. (2022). False memories in children's working memory. Retrieved from <https://osf.io/9gfmd/>
- Stahl, C., & Klauer, K.C. (2008). A simplified conjoint recognition paradigm for the measurement of gist and verbatim memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(3), 570. <https://doi.org/10.1037/0278-7393.34.3.570>
- Sugrue, K., & Hayne, H. (2006). False memories produced by children and adults in the DRM paradigm. *Applied Cognitive Psychology, 20*, 625–631. <https://doi.org/10.1002/acp.1214>
- Tam, H., Jarrold, C., Baddeley, A.D., & Sabatos-DeVito, M. (2010). The development of memory maintenance: Children's use of phonological rehearsal and attentional refreshment in working memory tasks. *Journal of Experimental Child Psychology, 107*, 306–324. <https://doi.org/10.1016/j.jecp.2010.05.006>
- Uittenhove, K., Chaabi, L., Camos, V., & Barrouillet, P. (2019). Is working memory storage intrinsically domain-specific? *Journal of Experimental Psychology: General, 148*(11), 2027-2057. <https://doi.org/10.1037/xge0000566>
- Wimmer, M. C., & Howe, M. L. (2009). The development of automatic associative processes and children's false memories. *Journal of Experimental Child Psychology, 104*(4), 447-465.