



## OPEN ACCESS

## EDITED BY

Mei-Ching Lien,  
Oregon State University, United States

## REVIEWED BY

James Robert Houston,  
Middle Tennessee State University,  
United States  
Davide Nardo,  
School of Clinical Medicine, University of  
Cambridge, United Kingdom

## \*CORRESPONDENCE

Michèle C. Muhmenthaler  
✉ michele.muhmenthaler@unibe.ch

## SPECIALTY SECTION

This article was submitted to  
Attention,  
a section of the journal  
Frontiers in Cognition

RECEIVED 16 December 2022

ACCEPTED 03 March 2023

PUBLISHED 23 March 2023

## CITATION

Muhmenthaler MC, Dubravac M and Meier B  
(2023) How attention and knowledge modulate  
memory: The differential impact of cognitive  
conflicts on subsequent memory—A review of  
a decade of research. *Front. Cognit.* 2:1125700.  
doi: 10.3389/fcogn.2023.1125700

## COPYRIGHT

© 2023 Muhmenthaler, Dubravac and Meier.  
This is an open-access article distributed under  
the terms of the [Creative Commons Attribution  
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# How attention and knowledge modulate memory: The differential impact of cognitive conflicts on subsequent memory—A review of a decade of research

Michèle C. Muhmenthaler<sup>1\*</sup>, Mirela Dubravac<sup>2</sup> and Beat Meier<sup>1</sup>

<sup>1</sup>Institute of Psychology, University of Bern, Bern, Switzerland, <sup>2</sup>Department of Psychological and Brain Sciences, Texas A&M University, College Station, TX, United States

In order to cope with cognitive conflicts, attention and knowledge are required. In some conditions, cognitive conflicts can boost subsequent memory and in other conditions, they can attenuate subsequent memory. The goal of the present study is to provide a narrative review of studies from the last decade in which Stroop or flanker conflicts, task switching, perceptual disfluency or semantic incongruence were manipulated at study. We propose an integrative framework considering attentional mechanisms and knowledge structures. Attentional mechanisms can refer to conflict resolution, which is required to explain the memory benefit for incongruent stimuli in Stroop and Flanker paradigms. Attentional mechanisms can also refer to attention allocation, which is required to explain the memory cost for targets and the memory benefit for task-irrelevant distractors in task-switching paradigms. Moreover, attention allocation policies can also account for the inconsistent results for perceptual disfluency manipulations. Prior knowledge is required to explain effects of semantic congruency and incongruency: Information that is expected, or congruent with prior knowledge, is better remembered, namely by pre-existing schemata. Moreover, information that is unexpected or incongruent with prior knowledge attracts attention and is better remembered. The impact of prior knowledge on memory performance thus results in a U-shape function. We integrate the findings according to this framework and suggest directions for future research.

## KEYWORDS

cognitive conflict, attention allocation, consolidation theory, schema-congruence theory, conflict-monitoring hypothesis, load theory of selective attention

## 1. Introduction

From the moment we wake up, we employ attention and prior knowledge to reach goals and navigate successfully through the day. For every task like getting dressed, making coffee, and driving to work, we focus on task-relevant information while ignoring task-irrelevant information. Different task-sets can sometimes conflict when we need to do two tasks at the same time (e.g., getting dressed and monitoring the time) or when we switch tasks. Cognitive conflicts can also arise when a situation does not match with our prior knowledge (e.g., handling a new coffee machine), or with our expectancies (e.g., the postman bringing our

parcels at a specific time) or when we face incongruent information (e.g., green traffic light signals us to drive but a pedestrian is crossing the street). In the laboratory, cognitive conflicts are induced with dual-task paradigms, task-switching paradigms, violation of expectancies (prediction errors), perceptual disfluency or incongruent trials in conflict paradigms as the Stroop or the Flanker task. During task performance, all these conflicts slow us down compared to the condition without (Rogers and Monsell, 1995; Wylie and Allport, 2000; Egner and Hirsch, 2005; Bugg, 2008; Kalanthroff and Henik, 2014). Although the effects on immediate task performance are similar, the consequences on long-term memory vary substantially. The aim of this review is to examine the underlying cognitive mechanism promoting memory costs and gains.

We review research from the past decade on the effects of different encoding manipulations involving conflict on subsequent memory. For inclusion, a basic requirement was that a study included trial-unique stimuli and a kind of conflict in the study phase and that memory for these stimuli was measured in a subsequent test phase. We integrate these findings in a framework that consists of two overarching main factors—*attentional mechanisms* and *knowledge structures*. Attentional mechanisms can explain memory effects produced by dividing attention, task switching, Stroop and Flanker like conflicts, the attentional boost effect and perceptual disfluency. This part of our framework is mainly based on the load theory of selective attention, explaining memory costs (Lavie, 2005, 2010), enriched with attentional enhancements, explaining memory benefits (Diemand-Yauman et al., 2011; Swallow and Jiang, 2013; Krebs et al., 2015; Ptok et al., 2021; LaPointe et al., 2022). Knowledge structures can explain memory effects produced by prior knowledge (i.e., schemata), novelty and prediction errors. This part of our framework is based on schema theory of memory and consolidation theories (von Restorff, 1933; Wang and Morris, 2010; Van Kesteren et al., 2012; Gilboa and Marlatte, 2017). Accordingly, memory performance follows a U-shape function with information that is congruent with prior knowledge and information that is novel at the endpoints of a continuum (Greve et al., 2019; Quent et al., 2021).

In 2012 the first study was published in the line of research which is the core of our review. In their seminal study, Richter and Yeung (2012) investigated how cognitive control influences memory encoding by applying a task-switching procedure in the study phase. In 2015, two studies investigated the impact of Stroop like conflicts on subsequent memory performance (Krebs et al., 2015; Rosner et al., 2015a). As these studies found opposing effects on memory, the debate about why *more* cognitive control can result in both, memory gains and memory losses, was launched and inspired multiple follow-up studies. Thus, our review focuses primarily on relevant behavioral studies from the past decade.

The literature of the other paradigms is somewhat older. The research on divided attention started in the 90's and originated from the episodic and declarative memory research (Tulving, 1985; Gardiner and Parkin, 1990; Craik et al., 1996; Naveh-Benjamin et al., 1998; Yonelinas, 2002). The studies mainly focused on the comparison of encoding and retrieval effects, and on different types of memory. The research on knowledge structures has some

origins in the educational and learning literature. Exploring which learning condition leads to the most efficient way of acquiring and consolidating new information is an important research topic since decades (Craik and Lockhart, 1972; Slamecka and Graf, 1978). This research also focused on pre-existing knowledge (i.e., schemata; Bjork and Allen, 1970; DeWinstanley and Bjork, 2004; Cepeda et al., 2006; Roediger and Karpicke, 2006; Rohrer and Taylor, 2007). Interestingly, the literature on perceptual disfluency also stems from the educational literature, as some researchers investigated the hypothesis that making study materials perceptually more difficult can promote sustainable learning (Hirshman et al., 1994; Sungkhasettee et al., 2011). The effects stemming from challenging encoding conditions were subsumed under the label “desirable difficulties” (Bjork and Bjork, 2011, 2020), emphasizing that more effort at study leads to better learning and memory (Craik, 2002; Staresina and Davachi, 2006; Staresina et al., 2009; Bjork and Kroll, 2015). The reviewed literature on knowledge structures moreover has its roots in neuroscientific and animal research on memory formation, consolidation and reconsolidation (Moscovitch et al., 2005; Tse et al., 2007; Wang and Morris, 2010; McClelland, 2013).

We present a narrative review of several lines of rather heterogenous research which we integrated in a framework. However, the main focus of our review remains on the task-switching and conflict studies which directly tested the implications of the load theory of selective attention and the conflict-monitoring hypothesis (Lavie and Cox, 1997; Botvinick et al., 2001, 2004; Lavie, 2005). The other paradigms have been reviewed elsewhere (Bennett and Flach, 1992; Wang and Morris, 2010; Van Kesteren et al., 2012; Swallow and Jiang, 2013; Weissgerber and Reinhard, 2017). An overview of the reviewed studies is presented in a [Supplementary Table 1](#).

In the first part, we provide theoretical considerations associated with attentional mechanisms and we include studies that used different paradigms to induce cognitive conflicts. [Figure 1](#) depicts an overview of different conflict paradigms. We use the term conflict for all types of encoding conditions which impair immediate task performance, such as dividing attention, task-switching, attentional boost, Stroop and Flanker conflicts and perceptual disfluency. In the second part, we provide theoretical considerations associated with knowledge structures and review research on the impact of prior knowledge and information that is incongruent with pre-existing knowledge and thus unexpected. We include studies that manipulated schema-congruence, novelty or induced prediction error. Our main goal was to review the relevant behavioral literature to provide an integrative view of the findings. As mentioned above, we started with the task switching and Stroop conflict studies, but then realized that the related paradigms need to be considered for the sake of completeness.

## 2. Attentional mechanisms

Allocating attention influences what we remember, and what we remember guides our attention (Becker and Rasmussen, 2008; Chun and Johnson, 2011). Attention and memory are therefore inextricably intertwined. Interactions between the two became the focus of research in the last decade and this field of research is



growing quickly. We will start with theoretical considerations about attentional mechanisms with a focus on the load theory of selective attention as this theory proved especially suitable for explaining the effects of a variety of encoding manipulations on memory (Lavie, 2005, 2010).

A basic premise of many attention theories is that attention is limited in capacity (Driver, 2001; Oberauer, 2019). In order to attain one's goals through focused and goal-oriented behavior, attention must be selective (Otten et al., 2002). Combining early and late selection processes, the theory of selective attention states

that perceptual and cognitive processing demands determine the selectivity of attention (Lavie, 2005; Tsal and Benoni, 2010). An intriguing assertion of this theory is that perceptual processing and cognitive processing both have capacity limits but exhausting these capacities would lead to opposing effects on selective attention. As perceptual processing operates automatically and perceptual resources are used obligatorily, task-irrelevant distractors are processed automatically when perceptual load is low. When perceptual load is high, however, the perceptual processing capacity is already exhausted by processing task-relevant information,

thus leading to higher selective attention toward targets (Lavie, 2005; Swallow and Jiang, 2013). Thus, perceptual load—defined by stimulus set characteristics (number, similarity, and type of presented items) and stimulus quality (e.g., readability, size, and color; Lavie et al., 2009)—*enhances* selective attention. When perceptual load is low, a second, higher-order control mechanism that actively inhibits attention to irrelevant distractors comes into play (Lavie, 2010). The efficiency of this control mechanism depends on the level of load on cognitive control functions such as working memory (Lavie, 2010; Tsal and Benoni, 2010). When cognitive load is low, attention is focused on task-relevant information and task-irrelevant distractors can be inhibited, thus selective encoding is intact. When the cognitive load is high, however, control functions are already absorbed by the target task which enables distractor intrusions, resulting in “broad” attention. Thus, cognitive load—defined by the proportion of time during which a given activity absorbs control functions (Barrouillet et al., 2007)—*reduces* selective attention.

## 2.1. Perceptual load

In line with the load theory of selective attention, several studies showed that various manipulations of perceptual load in a target task affect the processing of distractors (Lavie, 1995; Lavie and Cox, 1997; Brand-D’Abrescia and Lavie, 2007; Forster and Lavie, 2009). The theory predicts that higher perceptual load reduces distractor processing and thus interference from conflicting distractor information (i.e., attention is more selective). Most relevant for the present review, studies showed corresponding effects on subsequent memory (Jenkins et al., 2005; Lavie et al., 2009; Greene et al., 2017; Nussenbaum et al., 2017).

For example, Greene et al. (2017) investigated the impact of perceptual load on memory by showing video clips and measuring eye movements. Based on the premise that perceptual load reduces memory accuracy, they investigated whether the memory impairments under high perceptual load resulted from inattention blindness or from failure to visually inspect stimuli (due to capacity limits; Lavie et al., 2009). Their participants viewed a video depicting a theft either under high or low perceptual load and then had to identify characters from the video in photographic line-ups. High perceptual load impaired participants’ ability to identify the peripheral character (witness) but not the central character (thief). There was no effect of perceptual load on number of ocular fixations on the witness, time to first fixation or total visit duration. The authors concluded that memory impairments under high perceptual load were due to attentional failures rather than differences in visual search. Thus, this study confirmed that the perceptual load is the key factor in determining the locus of selective attention. When capacity limits of perception are reached, distractors in the periphery are not encoded, resulting in selective attention and memory.

Nussenbaum et al. (2017) also investigated the impact of perceptual load by exploring the effects of distractor number and content on memory. The authors asked participants to identify a target image from among zero, one, or three distractor images and to categorize the target as “alive” or “not alive.” A subsequent

recognition memory test assessed memory for target and distractor images. Results of Experiment 1 showed that target memory was *worse* in the one-distractor condition (low perceptual load) than in the three-distractor condition (high perceptual load). In contrast, distractor memory was *better* in the one-distractor condition, suggesting that a higher perceptual load leads to more selective memories (better target and worse distractor memory). Experiment 2 extended these findings by showing that a single distractor hurts target encoding more than three distractors, especially when the response associated with the distractor conflicts with the response to the target. This suggests that distractor number and content matters: As the number of distractors increases, distractor interference decreases probably due to the higher perceptual load leading to more selective attention to the target and better inhibition of distractors. In other words, when perceptual load is low, cognitive load takes effect. In the next sections we will discuss the effects of cognitive load on memory in more detail.

## 2.2. Cognitive load

Cognitive load is high when demands on cognitive control processes increase, for example due to working memory load or dual-task interference (Oberauer, 2019). Studies using divided attention or task-switching paradigms showed corresponding effects on memory (Lavie, 2005, 2010; Swallow and Jiang, 2013; Dubravac and Meier, 2022). In the following we will review these studies.

### 2.2.1. Divided attention

Effects of divided attention during encoding on later memory performance are widely documented. It is well-established that engaging in two tasks simultaneously results in impaired long-term memory (Gardiner and Parkin, 1990; Craik et al., 1996, 2018; Naveh-Benjamin et al., 1998; Fernandes and Moscovitch, 2000; Greene and Naveh-Benjamin, 2022). In a typical divided-attention procedure, participants have to perform two discrimination tasks at the same time, often involving different modalities (e.g., Middlebrooks et al., 2017; Greene and Naveh-Benjamin, 2022). The main finding is that dividing attention results in costs on subsequent memory performance.

For example, in a study by Craik et al. (2018), participants had to perform a word/non-word discrimination task in the primary task. In the divided-attention condition, half of the participants had to perform either a visual or an auditory continuous choice reaction time task. Across five experiments, the results demonstrated that dividing attention resulted in lower memory than full attention. The authors reasoned that dividing attention during encoding affected multiple types of processing, resulting in an overall degraded memory trace, rather than interrupting any particular process. In other terms, keeping two concurrent tasks active required additional control processes, which impaired target memory (Lavie, 2010; Oberauer, 2019).

In a study by Uncapher and Rugg (2005), the mechanisms underlying dividing attention were investigated by manipulating secondary task difficulty. In the primary task, an animacy decision

to visually presented study words had to be performed, the second task was to perform either an easy or a hard auditory monitoring task. The authors demonstrated that memory was worse when the words were encoded under hard secondary condition compared than under easy secondary task condition. This finding further corroborates that enhancing the cognitive load (hard vs. easy secondary task) reduces subsequent target memory.

### 2.2.2. Task switching

The task-switching paradigm has been developed to address the mechanisms of cognitive control by comparing task switch and task repetition trials (Rogers and Monsell, 1995; Wylie and Allport, 2000). On switch trials, reaction times are usually slower and error rates higher. These switch costs are highly robust and reflect the cognitive load imposed by the requirements to update the new task set and to select the appropriate task (Kiesel et al., 2010).

To investigate the memory consequences of task switching, Muhmenthaler and Meier (2019) presented participants with trial unique stimuli which they had to classify according to one of two categorization tasks in alternating runs (ABBA). In Experiment 1, participants classified images of animals as a mammal or bird (task A) and images of objects as a musical instrument or kitchen utensil (task B). These stimuli were univalent because animals could only be classified according to task A and objects could only be classified according to task B. In Experiment 2, participants classified images along the dimensions size (task A) and animacy (task B), thus turning them into bivalent stimuli (see Figure 1A). In both experiments, a subsequent surprise memory test assessed participants memory of the stimuli. The results showed that memory was worse for stimuli presented on switch trials. Critically, the effect size was three times larger in Experiment 2 with bivalent stimuli. The results indicate that task-set reconfiguration (present in both experiments) and between-task interference (present in Experiment 2) both impair subsequent memory performance. Critically, increasing cognitive load by additional between-task interference increased the switch cost for target memory, in line with the load theory of selective attention (Lavie, 2010).

Richter and Yeung (2012) investigated the question whether reconfiguration or between-task interference were critical for subsequent memory effects. The authors suggested that reconfiguration would divert resources away from encoding, thereby resulting in a general memory decline for task-relevant and task-irrelevant information. However, between-task interference would result in lower memory for *task-relevant* items, but higher memory for *task-irrelevant* items, due to residual attention to the former task (Monsell and Driver, 2000; Yeung et al., 2006). In order to disentangle between these possibilities, the authors used compound stimuli which consisted of pictures and words and participants had to switch between classifying them. Thus, each trial consisted of task-relevant (target) and task-irrelevant (distractor) information. In line with the between-task interference account, task switching impaired memory for task-relevant information but improved memory for task-irrelevant information. The authors concluded that task switching reduced the ability to encode the targets selectively and to inhibit distractors, that is, task switching resulted in impaired *memory selectivity*.

Chiu and Egner (2016) further developed the idea that between-task interference was crucial for the memory benefit for distractors in switch trials. Toward this goal, they examined two distractor categories. In one group, participants switched between two classification tasks, the distractors were relevant in one task and irrelevant in the other task. In the other group, the distractors (objects in the background) were never task-relevant (thus they were unrelated background pictures). The results showed better memory for distractors that were relevant in one task on switch compared to repeat trials, aligning with the finding by Richter and Yeung (2012): When the task switches, attention is broad, resulting in distractor intrusion. When the task is repeated, attention can be focused on the targets, thus less distractor intrusion. In contrast, the condition with the *always* irrelevant distractors resulted in lower memory for the distractors in switch compared to repeat trials. The authors concluded that the higher cognitive load in switch trials reduced memory performance. As the distractors were never task-relevant, this decline in memory performance must be seen as a memory *cost* and not as a successful strategy to inhibit distractors.

Building on this work, Dubravac and Meier (2022) directly tested whether higher cognitive load would lead to less selective memories. They employed a similar procedure as Richter and Yeung (2012, 2015) with picture-word pairs as stimuli. According to the time-based resource sharing model (Barrouillet et al., 2007), cognitive load is higher when attention demanding activities co-occur in time and concurrently draw on limited cognitive control resources that are necessary for selective attention (Lavie, 2010). Following this definition, the authors varied the cognitive load by manipulating task predictability, preparation time, and stimulus presentation duration in cued (Exp. 1–3) and alternating runs task switching procedures (Exp. 4 and 5). In all experiments, task switching resulted in lower memory selectivity. Shorter preparation times, however, reduced memory selectivity only in the cued task switching paradigm (cf., Richter and Yeung, 2012; Exp. 1), but not in the alternating runs task switching paradigm. This result suggests that a cue triggers preparatory processes that—given a long enough cue-to-stimulus interval (i.e., preparation time)—alleviate cognitive load during the classification task as task-set reconfiguration processes were preponed. Shorter stimulus presentation durations also reduced memory selectivity. With shorter stimulus presentation durations, the stimuli had to be held active in working memory to solve the classification task, thus increasing cognitive load during the task. This effect occurred mostly on switch trials, when cognitive load was already high. This further corroborates that co-occurring cognitive load results in lower memory selectivity.

Together, the studies about divided attention and task switching showed that cognitive load plays an important role. In *low* cognitive load conditions, such as full-attention or task repetitions, attention can be focused on the targets, which leads to improved target memory and lower distractor intrusion. In contrast, in *high* cognitive load conditions, such as divided-attention or task switching, the control mechanism is absorbed by other control processes which leads to broader attention (Lavie, 2010; Dubravac and Meier, 2022). Consequently, target memory is reduced and irrelevant distractors intrude. Further studies on divided attention and task-switching are in line with this interpretation and support the cognitive load theory (Gardiner and Parkin, 1990; Craik et al.,

1996, 2018; Richter and Yeung, 2015; Brito et al., 2016; Dubravac and Meier, 2020; Muhmenthaler and Meier, 2021a, 2022; Greene and Naveh-Benjamin, 2022; Naveh-Benjamin et al., 2022).

## 2.3. Transient shifts of attention and cognitive load

### 2.3.1. Conflict stimuli

Another line of research investigated the impact of conflict stimuli on subsequent memory performance (Krebs et al., 2015; Rosner et al., 2015a; Jiménez et al., 2020; Muhmenthaler and Meier, 2021a,b). It was suggested that detecting conflicts can enhance target memory (Botvinick et al., 2001, 2004; Yeung et al., 2004; Carter and Van Veen, 2007). According to the conflict-monitoring hypothesis, detecting conflicts leads to a transient upregulation of selective attention in anticipation of the next trial, resulting in improved conflict resolution (Egner and Hirsch, 2005). In this line of research, conflict stimuli are defined as stimuli which involve simultaneously active, competing representations which point to different responses (Egner et al., 2007; Bugg, 2008). Conflicts usually slow down immediate task performance and increase the error rates.

According to the load theory, selective encoding should be optimal under low cognitive load (Lavie, 2010; Tsal and Benoni, 2010). As conflict resolution *increases* the demands on control processes, the load theory would predict *lower* memory performance for conflict stimuli, however the opposite is true. Several studies showed that due to an up-regulation of attention, conflict stimuli can enhance later memory performance (Krebs et al., 2015; Rosner et al., 2015a; Ptok et al., 2019; Davis et al., 2020; Muhmenthaler and Meier, 2021a,b). Importantly, this up-regulation is only possible when the processing demands are low, that is, under low cognitive load. When the processing demands increase, interference is stronger than the attentional enhancement, thereby eliminating any beneficial effects on memory. In the following paragraphs, we review studies which investigated conflict stimuli in the light of these considerations.

Krebs et al. (2015) investigated the impact of conflict stimuli on memory using a variant of a face-word Stroop task. In the study phase, the participants performed a gender discrimination task with male or female faces which were overlaid with the words *man*, *house*, or *woman*, thus congruent, neutral, and incongruent face-word stimuli were created (see Figure 1B). At study, the participants had to judge the gender of the face while ignoring the superimposed word. In the subsequent recognition memory test, faces from incongruent trials resulted in better memory performance than faces from congruent or neutral trials. The authors concluded that the results were in line with the conflict-monitoring hypothesis postulating that interfering information triggers a top-down reinforcement and enhances attention to the targets (Egner and Hirsch, 2005; Egner et al., 2007).

In a replication attempt, Jiménez et al. (2020) found no evidence for conflict-enhanced memory. However, they found a memory benefit for incongruent trials following incongruent trials, suggesting that conflict over two successive trials might be

necessary to boost target memory. In the studies by Krebs et al. (2015) and Jiménez et al. (2020), the same procedure was used, but only Krebs et al. (2015) found conflict-enhanced memory. A reason might be that the inter-stimulus interval was up to 7 s in Krebs et al. study, as the experiment was conducted in the scanner, and 1 s only in the Jiménez et al. (2020) study. In the latter study, conflict effects may have been eliminated due to reduced processing time, that is, due to a high cognitive load. The cognitive load of a task can be seen as a function of the proportion of time during which it captures attention, thus impeding other attention-demanding processes, such as conflict resolution (Barrouillet et al., 2007). If conflict resolution is hampered, no memory benefit of conflicts can be expected.

In a study by Rosner et al. (2015a), the participants had to read the red word in a pair of red and green spatially interleaved words. They were instructed to read the red word aloud and to ignore the green distractor. Half of the items were congruent (the interleaved words had the same identity), and the other half were incongruent (the interleaved words had different identities), thus the incongruent condition triggered a conflict. Following the reading phase, participants completed a surprise recognition memory test. The results showed better memory performance for incongruent trials. The authors interpreted their findings in terms of selective attention processes which improved memory for incongruent trials.

Davis et al. (2020) extended this line of research by investigating context effects in a series of experiments. Overall, they replicated the results by Rosner et al. (2015a) and moreover demonstrated that blocked lists of congruence led to stronger subsequent conflict effects than mixed lists. This is in line with the finding that the memory benefit for incongruent trials occurred only on successive incongruent trials (Jiménez et al., 2020). A noteworthy feature of the studies by Rosner et al. (2015a) and Davis et al. (2020) was that the cognitive load was low, as the task consisted of word reading, which is almost automatic (Walczyk, 2000). Moreover, the participants had plenty of time to complete the task. If cognitive resources would have been more depleted by the tasks, conflict resolution would have been hampered, thereby probably eliminating any memory effects.

Evidence for this claim stems from a recent study, in which the research by Rosner et al. (2015a) and Davis et al. (2020) was extended. In this study, the cognitive load was enhanced by combining the conflict stimuli with a task-switching procedure (Muhmenthaler and Meier, 2021a). Task switching enhances the cognitive load due to the required task-set reconfiguration and the between-task interference (Rogers and Monsell, 1995; Wylie and Allport, 2000). The authors used similar stimulus materials as Rosner et al. (2015a), but instead of word reading, they applied a semantic word classification task which further enhanced the cognitive load. In the congruent condition, the two interleaved words derived from the same category whereas in the incongruent condition, the two words derived from different categories, thus they triggered a conflict. The results showed that the incongruent condition did not lead to better recall performance than the congruent condition. The authors then reduced the cognitive load in a follow-up experiment by presenting blocked instead of mixed lists (cf., Davis et al., 2020). The results revealed better memory for incongruent targets. Moreover, the results replicated the effect that

a series of incongruent trials might help to boost memory (Davis et al., 2020; Jiménez et al., 2020).

Ptok et al. (2019) investigated whether a memory benefit for conflict stimuli also occurs in a semantic priming procedure. In a series of experiments, they first showed a prime involving two words (e.g., “Kate / male”), then they showed a stimulus (e.g., “Kate”). The participants had to categorize the name by gender, then they assessed recognition memory for the stimuli. The authors varied the processing stages of the conflict and investigated whether a stage-specific focus of control demands would influence later memory. A memory benefit only occurred when the conflict lied on the semantic categorization stage (e.g., by showing “Kate / male” as prime) and not on the response stage (e.g., by showing “Kate / right”). They concluded that a memory benefit can be produced by semantic incongruency priming, but only when the primes induce a conflict at the semantic categorization stage. Moreover, in line with our account, they suggested that conflict-enhanced memory can only be observed when demands of the task are low and allow a degree of automaticity in responding.

In a recent study, the memory-enhancing effect was demonstrated by using a Flanker like conflict at study (Muhmenthaler and Meier, 2021b). In the congruent condition, participants saw three pictures, a target in the middle, and two identical flankers. Importantly, all the pictures were from the same stimulus category (e.g., three mammals). In the incongruent condition, participants saw also three pictures, but the target was from one category and two identical flankers were from another category (e.g., a mammal in the middle and two birds, see Figure 1C). Participants had to categorize the target and ignore the flankers. Different pictures in both conditions were used to eliminate potential effects of fluency or perceptual load, as in this setting, the perceptual difficulty was similar in both conditions. The results showed that the incongruent condition led to better target memory than the congruent condition, indicating that the memory-enhancing effect generalized from the Stroop to the Flanker task.

Together, the results of these studies show that cognitive conflicts such as Stroop or Flanker conflict can enhance subsequent memory performance for targets. When conflict is detected, attention is transiently enhanced, rather than reduced, as the load theory of attention would predict (Botvinick et al., 2004; Olivers and Meeter, 2008; Lavie, 2010). However, the attention enhancement is only viable under low cognitive load, that is, when processing of the target task is within the capacity limits of attention. In the next section we review the literature on the attentional boost effect. Similar to Stroop and Flanker like conflicts, we propose that the emergence of the attentional boost effect strongly depends on the cognitive load imposed by the target tasks.

### 2.3.2. The attentional boost effect

Due to the limited capacity of attention, the load theory of attention suggests that increasing attention to one task should reduce attention to another task, as in typical divided attention studies (Naveh-Benjamin et al., 1998; Lavie, 2010; Craik et al., 2018; Greene and Naveh-Benjamin, 2022). However, several studies showed that attending to a relevant target can actually boost the

perceptual processing of concurrent, but unrelated information, referred to as the attentional boost effect (Swallow and Jiang, 2010, 2013). It has been proposed that detecting a target produces a transient up-regulation of attention which enhances memory for items in close spatiotemporal proximity to the target (LaPointe et al., 2022).

In a seminal study, Swallow and Jiang (2010) asked participants to perform two continuous tasks at the same time. For one task the participants saw a series of scenes, one at a time at the center of the screen. Participants had to encode the scenes for a subsequent memory test. For a second task a stream of squares appeared superimposed over the scenes (see Figure 1D). The participants pressed a key as quickly as possible whenever a specified target square appeared. The square was completely unrelated to the scene. Then, a recognition test assessed memory for the scenes. The results showed better memory for the scenes presented with a target square than those presented with a distractor square. The authors concluded that increasing attention to a target can lead to widespread increases in perceptual processing, which enhances memory for them.

In Experiment 2 of the same study, the authors showed that auditory targets also facilitated image-encoding, thereby demonstrating that the attentional boost effect is not modality specific. In Experiment 3, the participants were instructed to memorize the scenes and to *ignore* the squares (single-task condition). In this experiment, the results showed no attentional boost effect and demonstrated that the effect depended on performing the target-detection task. Together, the authors concluded that the attentional boost effect might reflect the opening of an attentional gate, which enhances perceptual processing. In other words, target detection leads to a transient shift of attention, which later enhances memory for the unrelated scenes.

Although the occurrence of an attentional boost effect is widely documented (Swallow and Jiang, 2013; Mulligan et al., 2014; LaPointe et al., 2022), the exact underlying mechanism is not clear. As in classic divided attention studies, two tasks are performed simultaneously, but the outcomes are different. Like the conflict studies, the outcomes may depend on the processing demands associated with the two tasks, that is, on the cognitive load. In typical divided attention experiments, the tasks are rather complex, for example engaging in two discrimination tasks in different modalities (cf., Naveh-Benjamin et al., 1998; Craik et al., 2018; Greene and Naveh-Benjamin, 2022). The cognitive load is high, and up-regulations of attention are not viable. In contrast, in typical attentional boost experiments, the tasks are very simple, for example pressing the space button when a specific square appears. Processing is within the capacity limits of attention, as the tasks do not involve any higher-order control processes. The cognitive load is low and up-regulations of attention are viable (Oberauer, 2019).

This assumption is corroborated by the finding that the *single-task condition* of Experiment 3 (without targets) of the above-mentioned study led overall to better memory than the *dual-task condition* of Experiment 1 (Swallow and Jiang, 2010). Specifically, memory performance of the single-task condition was at the same level as the *peak* (when a target was detected) of the dual-task condition. This implies that overall, the dual-task condition *impaired* memory performance compared to the single-task condition, aligning with other dividing attention studies

which consistently showed memory costs (Craig et al., 2018; Greene and Naveh-Benjamin, 2022). In other terms, the attentional boost lifts performance up to the level of full attention performance but not beyond (LaPointe et al., 2022). However, as the cognitive load is low in attentional boost studies, it does not exceed working memory limits, and thus target detection can lead to trial-to-trial attentional enhancements.

More evidence for this claim stems from a further experiment of the study mentioned above (Swallow and Jiang, 2010; Experiment 5). In this experiment, the target squares could be green or red. In the *simple-detection* condition, participants pressed the spacebar whenever *either* a red or a green square appeared. In the *arbitrary-mapping* condition, participants pressed one key for red squares and another key for green squares. In the simple-detection task, an attentional boost effect occurred. Critically, in the arbitrary-mapping condition, no attentional boost effect emerged. It was eliminated due to increased processing demands associated with response selecting, that is, due to high cognitive load.

### 2.3.3. Perceptual disfluency

Studies of perceptual disfluency on memory have their roots in the desirable difficulties account. This account suggests that making things harder to learn can *improve* subsequent learning and memory (Bjork and Bjork, 2011; Maddox and Balota, 2015; Bjork and Yue, 2016). Several studies provided evidence that perceptual disfluency, despite slowing down immediate task performance, can improve subsequent memory performance. Similar to conflict and attentional boost studies, it has been proposed that disfluency results in a transient up-regulation of attention, in order to decode the stimulus and to optimize performance (LaPointe et al., 2022). However, the impact of perceptual disfluency on memory is more inconsistent, with some studies leading to better, other to worse memory and other showing no effects. Similar to the attentional boost effect and the conflict literature, cognitive load could moderate the relationship between perceptual disfluency and memory. When cognitive load is low, an up-regulation of attention is possible, and disfluency can act as a desirable difficulty (LaPointe et al., 2022). In contrast, when cognitive load is high, there are not enough resources available to decode the disfluent stimuli, and disfluency reduces subsequent memory (Eitel et al., 2014; Meier and Muhmenthaler, 2021). In the next paragraphs, we review the literature on perceptual disfluency in the light of these considerations.

In a study by Sungkhasettee et al. (2011) the participants had to study inverted and upright words. Participants were instructed that they had to read each word and to encode the words for a later test. The authors found that inverted words were more often recalled than upright words and interpreted the results in terms of processing fluency and desirable difficulties. The cognitive load imposed by word reading was low and almost automatic, thus perceptual disfluency enhanced later memory (Walczyk, 2000).

Diemand-Yauman et al. (2011) investigated whether disfluent fonts can lead to better memory than fluent fonts. In two experiments, the participants were asked to learn facts about three species of aliens, they had 90 s to memorize 21 features. The authors

found that that information presented in hard-to-read fonts was better remembered than information which was presented in easy-to-read fonts. They interpreted that hard-to-read fonts can operate as a desirable difficulty, which engendered deeper processing strategies (Bjork and Bjork, 2011). We assume that due to enough processing time, perceptual disfluency enhanced later memory.

Further studies showed that the memory-enhancing effect of disfluent fonts is not as robust as the study by Diemand-Yauman et al. (2011) implied. Seufert et al. (2017) provided evidence for a boundary condition associated with perceptual disfluency: They manipulated the disfluency level of the fonts and the results showed significant differences, indicating that there is an optimal level of disfluency on performance that leads to increased engagement. These results provided evidence that there is a breaking point of disfluency where the perceptual load begins to be too high, and disfluency begins to be an “undesirable difficulty.”

Meier and Muhmenthaler (2021) provided evidence for a reversed effect of disfluent fonts on subsequent recognition memory. The authors presented different statements of the type “a *City* is in a *Country*” in fluent (e.g., **Arial**) or disfluent fonts (e.g., **Harrington**), and the participants had to decide whether these statements were true or not as fast and as accurate as possible (see Figure 1E). A following recognition test revealed that fluent statements were better remembered than disfluent statements. Due to high cognitive load in the study phase (deciding whether a geographic statement was true and selecting the appropriate response under time pressure), there were not enough working memory resources available to decode the disfluent statements, thus disfluency acted as an undesirable difficulty which reduced subsequent memory.

Rosner et al. (2015b) investigated the impact of clear and blurred words on recognition memory performance (see Figure 1F). The participants had to read the words aloud in their own pace. In several experiments, the results showed that blurred words were better remembered than clear words. The authors concluded that the blurred words up-regulated cognitive control, which enhanced encoding and later memory. This is in line with our account, as the cognitive load imposed by word reading was low and enough working memory resources were left to decode the blurred words.

A somehow unexpected result was reported by Yue et al. (2013). The authors investigated the impact of blurred and clear words on metacognitive predictions and subsequent memory. In five experiments, the participants had to read blurred or clear words and they had to give a judgement of learning (JOL) after each word. The participants had plenty of time for encoding the words. Thus, due to low cognitive load, we would expect an advantage for the blurred words. However, the opposite was true, the authors reported overall a small memory benefit for clear words. The JOLs were higher for the perceptually fluent items, as more easily processed information is usually predicted to be more retrievable in the future (Hirshman et al., 1994; Rhodes and Castel, 2008). Giving JOLs activates pre-existing knowledge, or schemata (Staresina et al., 2009; Meier and Muhmenthaler, 2021). This mechanism probably enhanced memory for the clear words. Thus, it seems that the metacognitive judgements overwrote the subtle effects of perceptual disfluency, despite enough available

working memory resources. The study showed that small changes in the experimental designs can change or even reverse the memory effects, thus implying that it is noteworthy to thoroughly consider each manipulation.

In a recent study, LaPointe et al. (2022) combined two attentional manipulations that cause transient shifts of attention on memory, perceptual disfluency and attentional boost. They tested the hypothesis that the two attentional manipulations produce redundant effects on recognition. The participants had to read blurred or clear words while they had to monitor for and respond to target signals, whereas they had to ignore distractor signals. The results showed memory-enhancing effects for both, perceptual disfluency as well as for the target signals, however, the two factors did not interact. The authors concluded that the attentional effects for disfluency and target detection were additive and not redundant. Moreover, the authors reasoned that sufficient resources in response to target detection and to blurry words were available in this setting, which may have led to an additive memory effect of the two variables. It would be interesting to investigate these memory effects when the resources would approach or exhaust capacity limits. To investigate combined attentional manipulations at different levels of cognitive load may be an avenue for future research (cf., Muhmenthaler and Meier, 2021a).

Together, perceptual disfluency is basically associated with a memory enhancement, but the effect depends on the level of cognitive load, and it requires an optimal level of disfluency (Seufert et al., 2017). The disfluency effect is not as robust as for example the detrimental effects of task switching and dividing attention on subsequent recognition memory. For example, Eitel et al. (2014) as well as Rummer et al. (2016) reported null effects, despite applying the same experimental manipulation as Diemand-Yauman et al. (2011). Further research is necessary to explore the impact of other possible moderators in the relationship between perceptual disfluency and later memory.

## 2.4. Attentional mechanisms: Preliminary discussion

Because attention is limited in capacity, it must be selective in order to reach the most relevant goals (Driver, 2001; Oberauer, 2019). According to the load theory of attention, responding to a target should increase demands on control processes, thereby reducing memory. However, the literature on Stroop and Flanker like conflicts, the attentional boost effect and disfluency showed that responding to relevant targets can transiently enhance attention, rather than reduce it, as the load theory of attention would predict (Lavie, 2005; Olivers and Meeter, 2008). When specific targets appear, transient attentional enhancements are triggered, in order to optimize performance. Importantly, the cognitive load imposed by the target task can act as a moderator in this relationship. Under high cognitive load, the capacity limits of attention are reached, and up-regulations of attention are not viable. Under low load, however, trial-to-trial adaptations are viable, thereby leading to memory gains.

Noteworthy, the tale is not that simple and other factors may further modulate subsequent memory. For example, the study by Yue et al. (2013) showed that adding a simple action to the

experimental procedure can reverse the effects on subsequent memory. Giving a metacognitive judgment (i.e., a judgment of learning after each trial), reversed the memory effects in this study. Metacognition is based on prior knowledge, or schemata, which is the main topic of the next section. This results of this study foreshadows that the two components of our framework, attention and knowledge, can interact.

Of importance with regard to many studies in this field is the lack of baselines. For example, in the task-switching studies, we do not know whether task repeating enhances or decreases memory performance compared to a baseline condition. Memory performance resulting from *task-repeating* trials (trials from a task-switching block) should be compared with single-task trials (trials from a task block in which one task has to be performed repeatedly). Specifically, task-repeating trials compared to single-task trials could enhance memory due to a more cautious responding style, or due to higher motivation (Woodward et al., 2003). In contrast, task-repeating trials might decrease memory compared to single-task trials due to the higher cognitive load in these trials. Although not having direct evidence for this claim, we assume that the latter would apply. The lack of baselines is also an issue in the other conflict paradigms. The assessment of baselines should be considered in future research.

## 3. Knowledge structures

Our brains are optimized to remember relevant information and to quickly discard irrelevant information (Van Kesteren and Meeter, 2020). Our brains are also optimized to remember events that differ from previous experiences, as adaptation to the ever-changing world is critical for survival (Duszkiewicz et al., 2019). We preferentially attend to irregularities in the environment which may signal an upcoming danger (Reggev et al., 2018). Together, it is adaptive to remember information that is *congruent* as well as *incongruent* with prior knowledge, reflecting regularities and irregularities in the environment. It is therefore no surprise that these types of information are better remembered than neutral information.

### 3.1. Prior knowledge and novelty

Knowledge is represented in schemata which can be defined as networks of interconnected neocortical representations of prior knowledge. They are established in early childhood, continue to develop throughout life and allow us to make efficient judgements in an economical and adaptive way (Iran-Nejad and Winsler, 2000). Schemata exert powerful influences over how events are perceived and interpreted (Gilboa and Marlatte, 2017). They maximize the efficiency of new learning, expand memory capacity, and enable inferential processing (Bonasia et al., 2018). Events that are congruent with pre-existing schemata lead to better memory than incongruent events and this effect is stronger after a delay (Hennies et al., 2016). The *schema theory on memory* predicts that schemata enhance memory due to efficient encoding and accelerated consolidation processes (Tse et al., 2007; Wang and Morris, 2010; Van Kesteren et al., 2012; McClelland, 2013).

Novelty is also known to enhance memory (von Restorff, 1933; Bonasia et al., 2018), with the distinctiveness of a stimulus thought to improve later memory. Memory for a special event (e.g., the first kiss) can be much better than memory for events that occur repeatedly (Poppenk et al., 2010). Novelty involves responding to information that is not expected or predicted in a given context on the basis of prior experience (Van Kesteren et al., 2012; Quent et al., 2021).

According to these considerations, memory performance follows a *U*-shape, but the two ends are associated with different expressions of memory. Schema-congruence produces generalized semantic memories, reflecting schemata, whereas incongruence/novelty produces detailed episodic memory, reflecting the encoding of a “snapshot” (Quent et al., 2021). Both types of memory are mediated by different brain structures (Van Kesteren et al., 2012). In the next section, we review articles which investigated the impact of schema-congruence and novelty on subsequent memory performance.

In a study by Cortese et al. (2019), congruence effects for color-word associates were investigated in a Stroop color naming task. The participants had either to name the color of a font, or to read the word which was presented in a specific color. The words appeared in a congruent (e.g., *ocean* in *blue*), neutral (e.g., *lawyer* in *green*), or incongruent (e.g., *banana* in *blue*) manner. Then, the participants had to recall the words. The results of the memory test revealed that words which have been shown in the congruent condition were more often recalled than words from the incongruent which were more often recalled than the neutral condition. This pattern of results appeared in both tasks, font color naming and word reading. Together, the results showed the expected *U*-shape function of memory, but they also revealed that the effect for schema-congruence was stronger than the effect for novelty.

The congruence effect was further investigated by Van Kesteren M. T. R. et al. (2013). The authors explored the impact of subjective congruence involving different modalities. The participants had to study simultaneously presented combinations of visual motifs, visual object words and tactile fabric samples which were either congruent (e.g., jacket and leather) or incongruent (e.g., umbrella and lace). A recognition memory test for the motifs was administered either immediately, after 1 day, or after 2 days. The results showed that congruent stimuli were better remembered after a consolidation interval, but not in the immediate test. These results demonstrated that a consolidation phase may be a precondition for schema-congruence effects.

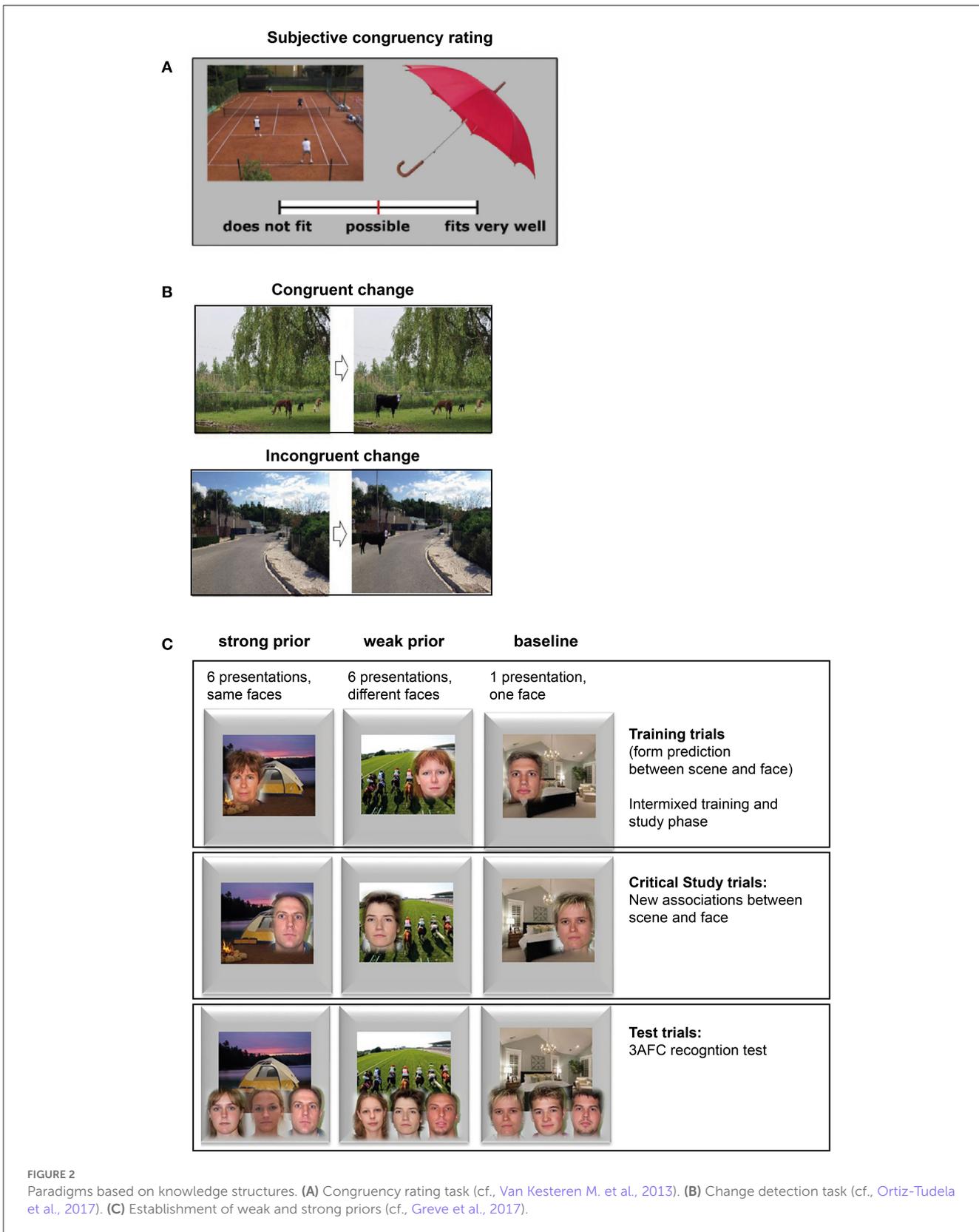
Van Kesteren M. et al. (2013) also investigated the impact of subjective schema-congruence on later recognition memory. They used pairs of objects and scenes at encoding, and the participants indicated how congruent they found these pairs and were tested on recognition memory for these associations 1 day later. For example, a congruent pair was *tennis court—tennis racket* and an incongruent pair was *classroom—soup ladle* (see Figure 2A). The authors found a monotonic increase in memory performance with increasing congruency ratings. They did not find evidence of enhanced memory for incongruent information. The authors interpreted the results as confirmation of the schema theory of memory (Wang and Morris, 2010; McClelland, 2013).

In the studies mentioned above, the schemata reflected prior knowledge of the participants. Hennies et al. (2016) investigated whether a schema can be induced *experimentally*, through several training sessions. The participants learned facts over six sessions during 2 weeks. Then, they learned new facts which were related or completely unrelated to the schema they had established. Directly after these sessions, memory for all facts was tested in a two-alternative task. The results revealed better memory for schema-congruent information, thus providing evidence that a schema can be established within a few days, and that this schema can lead to memory gains.

Ortiz-Tudela et al. (2017) used a change-detection task to explore long-term consequences of schema congruence. They manipulated the congruence between a changing object and a background scene (see Figure 2B). For example, they showed a cow in the prairie (congruent with prior knowledge) or a cow on the street (incongruent with prior knowledge), the background-only and the background-plus-target images were presented in rapidly alternating sequences to generate a flickering appearance. The participants had to press a specific button when they detected an object in the foreground. Across three experiments, the data showed that incongruent events were faster detected than congruent events. However, the results of the memory test revealed that schema-congruent events led to better recognition memory performance than incongruent events, providing evidence that the schema effect generalizes to other tasks, such as the change-detection task.

Bonasia et al. (2018) let their participants watch narrative film clips which contained events that were either congruent with prior knowledge or not. Memory for the events was tested either immediately or after one week. Both congruence with prior knowledge and incongruence/novelty enhanced memory for events, though incongruent events were recalled with more errors over time. The authors concluded that novel and congruent information both enhanced memory but were processed *via* distinct mechanisms. The findings confirmed that memory performance was a *U*-shape function of congruence (Van Kesteren et al., 2012). The authors emphasized that they could demonstrate these effects with more naturalistic events than usually used in the labs, such as film clips.

Greve et al. (2019) investigated the impact of expectations (i.e., schema-congruent) and events that conflict with schemata (i.e., schema-incongruent) on subsequent memory. Across four experiments, schemata were established by training relationships between randomly paired objects (e.g., a shoe and an umbrella). Thus, as in the Hennies et al. (2016) study, the schemata were induced experimentally and did not reflect prior knowledge. The participants learned which of two types of objects had a higher value (the rule). In congruent conditions, the rule remained constant across trials; in incongruent conditions, the rule reversed after the penultimate trial, that is, before the final critical trial that was later tested; in unrelated conditions, rules reversed several times. Thus, in congruent and incongruent conditions, the schema was either violated (incongruent condition) or not (congruent condition) on a critical trial. Better memory was found for both congruent and incongruent trials, relative to unrelated trials, producing memory performance that followed a *U*-shape function of congruence.



The congruence advantage but not incongruence advantage was mediated by post-encoding processes, whereas the incongruence advantage, but not congruence advantage, emerged even if the

incongruent information was irrelevant to the schema. The authors concluded that schemata augmented memory in multiple ways, depending on the match between novel and existing information.

Reggev et al. (2018) examined the role of experimental distinctiveness as a potential explanation for the memory benefits for novel and schema-congruent information. Across two experiments, they used word pairs which were either familiar (e.g., yellow banana) or novel (e.g., yellow zebra). The participants had to decide whether the word pairs were congruent or not. In a later test phase, recognition memory for the nouns was assessed. The results revealed that novelty was sensitive to its experimental proportions: improved memory for novelty was observed when novel word pairs were relatively rare. Memory levels for schema-congruent items, in contrast, were completely unaffected by experimental proportions, highlighting their insensitivity to list-based distinctiveness. The authors concluded that novel and congruent items both enhanced memory but were processed *via* partially distinct mechanisms.

The results of these studies provide evidence for a robust effect of schema-congruence on subsequent memory. The schema theory of memory explains this effect with an accelerated integration of new information into a pre-existing schema (Tse et al., 2007; Wang and Morris, 2010; McClelland, 2013; Durrant et al., 2015). Despite applying different tasks at study and applying different memory tests, the memory advantage for schema-congruent information materialized. The results also showed that a consolidation phase increases the effects, or may even be necessary for an effect to materialize (Van Kesteren M. T. R. et al., 2013; Hennis et al., 2016). The studies also provided evidence for a memory advantage for novel or unexpected events, but the results revealed that this effect is less robust than the schema effect on memory. The effect depends on the precise nature of the encoding and retrieval tasks, such as context effects (Reggev et al., 2018). In the next section we review research on the impact of prediction errors, which can be seen as a special case of novelty.

### 3.2. Prediction error

When our experience violates our predictions, it is adaptive to update our knowledge, in order to make better and more accurate predictions in the future (Bein et al., 2021). Theoretical models propose that such prediction errors should be encoded as distinct memory traces, reflecting the encoding of a “snapshot,” to prevent that previous memories interfere (McClelland, 2013; Quent et al., 2021). For example, the “Predictive Interactive Multiple Memory Signals” (PIMMS) framework is a framework for understanding how prior knowledge and prediction errors affect memory formation (Greve et al., 2017, 2019). According to this framework, the brain is assumed to contain hierarchical representations of the world, where representations at one level of the hierarchy predict the activity of representations in lower levels. The difference between those predictions and the sensory evidence from lower levels comprises the prediction error, which is assumed to drive learning between levels, so as to improve predictions and minimize prediction errors in the future (Friston, 2005). PIMMS offers a framework for considering how prediction errors might vary in the world, and therefore be manipulated experimentally in the laboratory.

Greve et al. (2017) examined the implications of the PIMMS framework. Specifically, the authors examined whether a prediction error reflects the divergence between the prior probability (from previous experiences) and sensory evidence (from the current experience). The hypothesis was tested across five experiments, in terms of peoples’ ability to encode a single presentation of a scene-item pairing as a function of previous exposures to that scene and/or item. Memory was tested by presenting the scene together with three choices for the previously paired item, in which the two foil items were from other pairings within the same condition as the target item. For example, the prior expectation was manipulated by training the participants to associate a scene with one or more unfamiliar faces (see Figure 2C). In order to induce a strong prior, a scene was repeatedly paired with the same face during training (high prediction error condition). In order to induce a weak prior, a scene was paired with different faces (low prediction error condition). Critical study trials used novel faces, evoking a higher degree of prediction error when the prior was strong. In a subsequent test phase, associative memory for faces paired in the critical study trials was assessed. The results showed, as hypothesized, better memory for the new scene-face pairing in the high prediction error condition compared to the low prediction error condition, that is, memory was best when the expectation was violated. The prediction error hypotheses were supported in all experiments. The prediction errors reflected the divergence between the prior probability and current sensory evidence, and the authors concluded that the PIMMS framework provided the most parsimonious account of the pattern of results.

In a study by Quent et al. (2021), the relationship between object-location memory and expectedness was investigated across four experiments. In an immersive virtual reality, participants explored a virtual kitchen with the instruction to count and memorize the locations of 20 objects that were scattered across the room. In the following recall phase, the participants reentered the kitchen (now without the 20 original objects), were given one object, and were asked to place it at its previously seen location. Once placed, the object disappeared, and the process was repeated for the remaining 19 objects. Recall was followed by a recognition test, performed outside the virtual reality. Each trial showed one studied object in three alternative locations, one of which was correct. Importantly, the target and two foil locations were matched in expectancy according to the normative ratings, so using prior knowledge to guess the location could not help performance. The results demonstrated better memory for highly expected and unexpected locations relative to neutral locations. The results showed that memory followed a U-shaped function of the expectancy of an event, with better memory for highly expected or highly unexpected object locations.

In a study by Bein et al. (2021), the authors investigated the impact of experience violations on memory, by repeatedly exposing participants to pairs of objects. During a prediction learning phase, the participants were presented with a stream of objects that included neighboring pairs of objects that followed each other back to back, thereby evoking predictions. Then, they violated this prediction in half of these pairs during the critical violation phase by replacing the second object in the pairs with a novel object. The following item memory test required participants to

discriminate between identical old items and similar lures. The results revealed that the prediction errors enhanced recognition memory: Participants correctly identified more old items as old when those items violated expectations during learning, compared with items that did not violate expectations. Importantly, this memory enhancement was only observed when participants later showed intact memory for the related pairs which were used to establish the predictions. The authors concluded that the advantage for prediction errors was dependent on the strength of the predictions, defined as the participants' memory of the original pair. In a follow-up experiment, the authors reduced prediction strength by lowering associative binding during encoding and found that the memory advantage for violations was diminished.

Ortiz-Tudela et al. (2018) provided empirical evidence for the claim of Bein et al. (2021), suggesting that strong predictions are a necessary precondition to facilitate memory. The authors explored expectation violations by means of a validity paradigm. Across seven experiments, the authors showed participants arrows which pointed to specific directions. The participants were told that the relevant arrow (presented in a specific color) would point to the location at which the upcoming word stimulus was most likely to appear. On most of the trials, the stimulus appeared at the predicted location (expectation match), and on several trials, the stimulus appeared at another location (expectation mismatch). The authors reported evidence for a null effect of expectation violations on memory formation. We assume that using this task, the predictions were not strong enough to facilitate memory for violations (Bein et al., 2021). The spatial task might not have produced strong predictions, or priors, as the arrows may rather have been used as spatial cues. In other terms, the expectations were not based on previous memories, which may be a precondition for a memory gain.

### 3.3. Knowledge structures: Preliminary discussion

Taken together, the studies on schema-congruence, novelty and prediction errors provide robust evidence for the hypothesis that memory performance follows a *U*-shape, with better memory for schema-congruence at one end and incongruence/novelty at the other end of a continuum. The quality of these memories is however different and mediated by different brain structures (Van Kesteren et al., 2012; Greve et al., 2017). Schema-congruence is associated with generalized semantic memory and less hippocampal activity, whereas incongruence/novelty is associated with detailed episodic memory and enhanced hippocampal activity (Quent et al., 2021). The *U*-shape reflects the opposing demands of benefitting from reoccurring regularities to enable efficient encoding on the one hand, and on the other hand, of accommodating surprising information and irregularities in the environment, which is essential for flexible adaptation to an ever-changing environment (Greve et al., 2019).

## 4. General discussion

In the present article, we reviewed the differential impact of cognitive conflicts on subsequent memory. We distinguished

between conflict conditions which required more vs. less cognitive control or attentional resources and conflict conditions related to the (in-)congruence of the study materials with existing schemata. Thus, we distinguished conflicts based on attentional mechanisms and conflicts based on knowledge structures.

For attentional conflicts, the reviewed studies suggest that conflicts associated with dividing attention and task switching hurt subsequent memory, due to enhanced cognitive load (Lavie, 2010). Performing two tasks simultaneously loads working memory and this leads to interference effects, which later results in reduced memory performance compared to a full-attention condition (Naveh-Benjamin et al., 1998; Greene and Naveh-Benjamin, 2022). Task switching also hurts subsequent target memory. Specifically, stimuli which have been shown in switch trials lead to worse memory than stimuli which have been shown in repeat trials, as the cognitive load is enhanced in switch trials due to higher processing demands (Muhmenthaler and Meier, 2019, 2022). The high cognitive load in switch trials leads to a "broad" attention, thereby leading to distractor intrusion (Lavie, 2005, 2010; Richter and Yeung, 2012, 2015; Chiu and Egner, 2016; Dubravac and Meier, 2022). Neuroimaging studies demonstrated that processing associated with task switching and divided attention are mostly associated to activity in the prefrontal regions and related networks in parietal regions, highly depending on specific features of the task at study (Reynolds et al., 2004; Johnson and Zatorre, 2006; Johnson et al., 2007; Niendam et al., 2012; Grange and Houghton, 2014).

In contrast, conflicts associated with Stroop or Flanker like conflicts, the attentional boost effect and perceptual disfluency can enhance subsequent memory (Swallow and Jiang, 2010; Diemand-Yauman et al., 2011; Sungkhasettee et al., 2011; Mulligan et al., 2014; Krebs et al., 2015; Muhmenthaler and Meier, 2021b). By going through the literature, the present review provided evidence that such transient trial-to-trial attentional shifts are however only viable under low cognitive load. Under high load, the capacity limits of attention are exhausted and attentional enhancements are not viable. Thus, for attentional conflicts, we identified the presence of specific target stimuli, combined with the level of the cognitive load, as the critical variables, that determine the direction of a specific effect for a particular conflict manipulation. According to the conflict-monitoring hypothesis, the dorsal anterior cingulate cortex signals the regulative components in the dorsolateral prefrontal cortex the detections of conflicts, thereby creating a feedback loop between the two components (Botvinick et al., 2004). A neuroimaging study showed that the memory benefit of incongruent trials of a Stroop like conflict was associated with activity in these structures, thereby providing neuronal evidence for the conflict-monitoring hypothesis (Krebs et al., 2015). Whether these structures are also involved in attentional boost and disfluency effects is not documented and may be a topic for future research.

For conflicts at the level of knowledge structures, a somewhat different pattern emerged. Both high congruence and high incongruence with prior knowledge can benefit memory, leading to a *U*-shape function, with better memory performance for information that is congruent with an existing schema, and better memory performance for novel or unexpected information at both ends. The quality of these memories is different. Schema-congruence is associated with generalized semantic memory whereas incongruence/novelty is associated with detailed episodic

memory (Quent et al., 2021). The quality of these memories is also mediated by different neuronal structures. The neuroscientific model “schema-linked interactions between medial prefrontal and medial temporal lobe” (SLIMM) proposes a time-dependent shift from medial temporal lobe to neocortical representations (Van Kesteren et al., 2012; Quent et al., 2021). Within SLIMM, the medial prefrontal cortex is to detect the congruency of new information with existing information in neocortex. Greater congruence leads to greater medial prefrontal cortex activity, which is assumed to potentiate direct connections between neocortical representations. In contrast, associations that are incongruent with a schema or highly novel are encoded in the medial temporal lobe, comprising hippocampus, perirhinal and entorhinal cortices, and parahippocampal gyrus (Van Kesteren et al., 2012; Quent et al., 2021).

The different patterns within each domain underline the usefulness of the conceptual distinction of our framework. The conceptual distinction is represented in the involvement of different neural substrates, as specified above. Overall, attentional conflicts activate areas in the prefrontal cortex and related networks (Johnson et al., 2007; Niendam et al., 2012). In conditions of conflict resolution, the anterior cingulate cortex has been located as critical structure which signals the dorsolateral prefrontal cortex to recruit attentional resources when conflicts are detected (Botvinick et al., 2004). In contrast, for conflicts at the level of knowledge structures, the medial temporal lobe with the hippocampus and the medial prefrontal cortex are the most relevant structures (Van Kesteren et al., 2012; Quent et al., 2021). Thus, the two parts of our framework, attention and knowledge, are not only distinct in terms of theories, paradigms and outcomes, but also in terms of the involved neuronal structures.

Recently, Craik and Bialystok (2006) proposed a similar distinction in a framework to address cognitive changes across the lifespan. Specifically, they distinguished between “cognitive representation” and “control” as factors that have different lifespan trajectories, with stronger decline across adult age in control than in representations. Combining the ideas from Craik and Bialystok with our distinction of cognitive conflict domains, an avenue for future research would be to test the hypothesis that conflicts at the level of attentional mechanisms would result in stronger age-effects than conflicts at the level of knowledge structures.

Another avenue for further research is to investigate the combined impact of attention and knowledge on subsequent memory. We are not aware of studies which manipulated both, prior knowledge and cognitive control demands, with the purpose to explore their interactions. In the attentional section, we emphasized the importance of cognitive load at study. However, we believe that cognitive load would not moderate memory effects resulting from knowledge or expectation manipulations, as these manipulations do not load on working memory, and thus they do not stress the limited attentional resources. As mentioned above, these manipulations are mediated by different neuronal structures (Van Kesteren et al., 2012; Quent et al., 2021). The same might apply for perceptual load. Thus, we would not expect interactions between cognitive/perceptual load and knowledge manipulations, but rather additive effects (e.g., processing schema-congruent information in a full-attention

condition could maximally improve memory). However, there is evidence that pre-existing schemata can overwrite the memory-enhancing effects of attentional manipulations. In the study by Yue et al. (2013), the pattern of results showed that higher metacognitive judgments for fluent words eliminated the expected memory benefit for disfluent words. The results imply that the effects of schema-congruence might be stronger than the effects produced by perceptual disfluency. It is also conceivable that in some conflict studies, schema congruence somehow counteracted the incongruence effects (e.g., a yellow banana is schema-congruent, but a blue banana is Stroop-incongruent, both promoting better memory), thereby maybe explaining the rather small memory effects produced by Stroop like conflicts (Jiménez et al., 2020; Muhmenthaler and Meier, 2021a). The example nicely shows that it is important to carefully design these experiments. Together, there is evidence that the two components, attention and knowledge, can however interact and further affect memory.

It can be considered as a limitation of the present review that we did not discuss the types of memory tests that were used. Rather, we focused on the encoding manipulations. We are aware that the specific test may influence the resulting memory effects. However, so far, the evidence indicates that the direction of memory effects is quite robust across different memory tests. For example, in the domain of task switching, different methods have resulted in convergent evidence (Richter and Yeung, 2012; Muhmenthaler and Meier, 2019, 2021a), as switch costs on memory generalized across recognition, confidence ratings and free recall as memory measures. Nevertheless, addressing the impact of cognitive conflict systematically across different memory tests may be an interesting avenue for future research (e.g., free recall vs. recognition; explicit vs. implicit memory, etc.).

Another limitation may be that we mainly focused on studies which used short study-test intervals, although we are aware that the study-interval may also play a crucial role for memory effects, due to post-encoding and consolidation processes (Wang and Morris, 2010; Van Kesteren et al., 2012; Durrant et al., 2015). The literature on schema-congruence showed that a consolidation phase may even be a precondition for some memory effects (Van Kesteren M. T. R. et al., 2013). Nevertheless, systematically addressing the impact of different retention intervals (and more generally, time for consolidation) may be a fruitful line for future research.

## 5. Conclusion

In this review, we distinguished between cognitive conflicts at the level of attentional mechanisms and at the level of knowledge structures. We illustrated that in both domains the specific encoding situation must be taken into account to foresee whether the particular conflict results in a loss or gain for subsequent memory performance. Besides of ordering these phenomena on a theoretical level, these insights may help to create study situations to optimize learning. Toward this goal it might be relevant to investigate the combined effect of beneficial cognitive conflicts identified in this review.

## Author contributions

MM, MD, and BM wrote the manuscript. All authors approved the final manuscript for submission.

## Funding

MD received financial support from the Swiss National Science Foundation (Grant No.: P2BEP1\_200018). Open access funding by University of Bern.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., and Camos, V. (2007). Time and cognitive load in working memory. *J. Exp. Psychol.* 33, 570–585. doi: 10.1037/0278-7393.33.3.570
- Becker, M. W., and Rasmussen, I. P. (2008). Guidance of attention to objects and locations by long-term memory of natural scenes. *J. Exp. Psychol.* 34, 1325–1338. doi: 10.1037/a0013650
- Bein, O., Plotkin, N. A., and Davachi, L. (2021). Mnemonic prediction errors promote detailed memories. *Learn. Mem.* 28, 422–434. doi: 10.1101/lm.053410.121
- Bennett, K. B., and Flach, J. M. (1992). Graphical displays: Implications for divided attention, focused attention, and problem solving. *Hum. Fact.* 34, 513–533. doi: 10.1177/001872089203400502
- Bjork, E. L., and Bjork, R. A. (2011). “Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning,” in *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*, eds M. A. Gernsbacher, R. W. Pew, L. M. Hough, and J. R. Pomerantz (New York, NY: Worth), 56–64.
- Bjork, R. A., and Allen, T. W. (1970). The spacing effect: Consolidation or differential encoding? *J. Verb. Learn. Verbal Behav.* 9, 567–572. doi: 10.1016/S0022-5371(70)80103-7
- Bjork, R. A., and Bjork, E. L. (2020). Desirable difficulties in theory and practice. *J. Appl. Res. Mem. Cogn.* 9, 475–479. doi: 10.1016/j.jarmac.2020.09.003
- Bjork, R. A., and Kroll, J. F. (2015). Desirable difficulties in vocabulary learning. *Am. J. Psychol.* 128, 241–252. doi: 10.5406/amerjpsyc.128.2.0241
- Bjork, R. A., and Yue, C. L. (2016). Commentary: Is disfluency desirable? *Metacogn. Learn.* 11, 133–137. doi: 10.1007/s11409-016-9156-8
- Bonasia, K., Sekeres, M. J., Gilboa, A., Grady, C. L., Winocur, G., and Moscovitch, M. (2018). Prior knowledge modulates the neural substrates of encoding and retrieving naturalistic events at short and long delays. *Neurobiol. Learn. Mem.* 153, 26–39. doi: 10.1016/j.nlm.2018.02.017
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624. doi: 10.1037/0033-295X.108.3.624
- Botvinick, M. M., Cohen, J. D., and Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends Cogn. Sci.* 8, 539–546. doi: 10.1016/j.tics.2004.10.003
- Brand-D’Abrescia, M., and Lavie, N. (2007). Distractor effects during processing of words under load. *Psychon. Bull. Rev.* 14, 1153–1157. doi: 10.3758/BF03193105
- Brito, N. H., Murphy, E. R., Vaidya, C., and Barr, R. (2016). Do bilingual advantages in attentional control influence memory encoding during a divided attention task? *Bilingualism* 19, 621–629. doi: 10.1017/S1366728915000851
- Bugg, J. M. (2008). Opposing influences on conflict-driven adaptation in the Eriksen flanker task. *Mem. Cogn.* 36, 1217–1227. doi: 10.3758/MC.36.7.1217
- Carter, C. S., and Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cogn. Affect. Behav. Neurosci.* 7, 367–379. doi: 10.3758/CABN.7.4.367
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., and Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychol. Bull.* 132, 354–380. doi: 10.1037/0033-2909.132.3.354
- Chiu, Y. C., and Egner, T. (2016). Distractor-relevance determines whether task-switching enhances or impairs distractor memory. *J. Exp. Psychol.* 42, 1. doi: 10.1037/xhp0000181
- Chun, M. M., and Johnson, M. K. (2011). Memory: Enduring traces of perceptual and reflective attention. *Neuron* 72, 520–535. doi: 10.1016/j.neuron.2011.10.026
- Cortese, M. J., Khanna, M. M., and Von Nordheim, D. (2019). Incidental memory for colour word associates processed in colour naming and reading aloud tasks: Is a blue ocean more memorable than a yellow one? *Memory* 27, 924–930. doi: 10.1080/09658211.2019.1607877
- Craik, F., and Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends Cogn. Sci.* 10, 131–138. doi: 10.1016/j.tics.2006.01.007
- Craik, F. I. M. (2002). Levels of processing: Past, present ... and future? *Memory* 10, 305–318. doi: 10.1080/09658210244000135
- Craik, F. I. M., Eftekhari, E., and Binns, M. A. (2018). Effects of divided attention at encoding and retrieval: Further data. *Mem. Cogn.* 46, 1263–1277. doi: 10.3758/s13421-018-0835-3
- Craik, F. I. M., and Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *J. Verb. Learn. Verbal Behav.* 11, 671–684. doi: 10.1016/S0022-5371(72)80001-X
- Craik, F. I. M., Naveh-Benjamin, M., Govoni, R., and Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *J. Exp. Psychol.* 125, 159–180. doi: 10.1037/0096-3445.125.2.159
- Davis, H., Rosner, T. M., D’Angelo, M. C., MacLellan, E., and Milliken, B. (2020). Selective attention effects on recognition: The roles of list context and perceptual difficulty. *Psychol. Res.* 84, 1249–1268. doi: 10.1007/s00426-019-01153-x
- DeWinstanley, P. A., and Bjork, E. L. (2004). Processing strategies and the generation effect: Implications for making a better reader. *Mem. Cogn.* 32, 945–955. doi: 10.3758/BF03196872
- Diemand-Yauman, C., Oppenheimer, D. M., and Vaughan, E. B. (2011). Fortune favors the bold (and the Italicized) Effects of disfluency on educational outcomes. *Cognition* 118, 111–115. doi: 10.1016/j.cognition.2010.09.012
- Driver, J. (2001). A selective review of selective attention research from the past century. *Br. J. Psychol.* 92, 53–78. doi: 10.1348/000712601162103
- Dubravac, M., and Meier, B. (2020). Stimulating the parietal cortex by transcranial direct current stimulation (tDCS): No effects on attention and memory. *AIMS Neuroscience* 8, 33–46. doi: 10.3934/Neuroscience.2021002
- Dubravac, M., and Meier, B. (2022). Cognitive load enhances memory selectivity. *Quart. J. Exp. Psychol.* 2022, 17470218221132846. doi: 10.1177/17470218221132846

## Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcogn.2023.1125700/full#supplementary-material>

- Durrant, S. J., Cairney, S. A., McDermott, C., and Lewis, P. A. (2015). Schema-conformant memories are preferentially consolidated during REM sleep. *Neurobiol. Learn. Mem.* 2, 11. doi: 10.1016/j.nlm.2015.02.011
- Duszkiewicz, A. J., McNamara, C. G., Takeuchi, T., and Genzel, L. (2019). Novelty and dopaminergic modulation of memory persistence: A tale of two systems. *Trend. Neurosci.* 42, 102–114. doi: 10.1016/j.tins.2018.10.002
- Egner, T., Delano, M., and Hirsch, J. (2007). Separate conflict-specific cognitive control mechanisms in the human brain. *NeuroImage* 35, 940–948. doi: 10.1016/j.neuroimage.2006.11.061
- Egner, T., and Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nat. Neurosci.* 8, 1784–1790. doi: 10.1038/nn1594
- Eitel, A., Kühl, T., Scheiter, K., and Gerjets, P. (2014). Disfluency meets cognitive load in multimedia learning: Does harder-to-read mean better-to-understand? *Appl. Cogn. Psychol.* 28, 488–501. doi: 10.1002/acp.3004
- Fernandes, M. A., and Moscovitch, M. (2000). Divided attention and memory: Evidence of substantial interference effects at retrieval and encoding memory-guided attention (contextual cueing) view project. *J. Exp. Psychol.* 2, 155. doi: 10.1037/0096-3445.129.2.155
- Forster, S., and Lavie, N. (2009). Harnessing the wandering mind: The role of perceptual load. *Cognition* 111, 345–355. doi: 10.1016/j.cognition.2009.02.006
- Friston, K. (2005). A theory of cortical responses. *Philos. Trans. Royal Soc. B.* 360, 815–836. doi: 10.1098/rstb.2005.1622
- Gardiner, J. M., and Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Mem. Cogn.* 18, 579–583. doi: 10.3758/BF03197100
- Gilboa, A., and Marlatte, H. (2017). Neurobiology of schemas and schema-mediated memory. *Trends Cogn. Sci.* 21, 618–631. doi: 10.1016/j.tics.2017.04.013
- Grange, J., and Houghton, G. (2014). *Task Switching and Cognitive Control*. Oxford University Press.
- Greene, C. M., Murphy, G., and Januszewski, J. (2017). Under high perceptual load, observers look but do not see. *Appl. Cogn. Psychol.* 31, 431–437. doi: 10.1002/acp.3335
- Greene, N. R., and Naveh-Benjamin, M. (2022). The effects of divided attention at encoding on specific and gist-based associative episodic memory. *Mem. Cogn.* 50, 59–76. doi: 10.3758/s13421-021-01196-9
- Greve, A., Cooper, E., Kaula, A., Anderson, M. C., and Henson, R. (2017). Does prediction error drive one-shot declarative learning? *J. Mem. Lang.* 94, 149–165. doi: 10.1016/j.jml.2016.11.001
- Greve, A., Cooper, E., Tibon, R., and Henson, R. N. (2019). Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. *J. Exp. Psychol.* 148, 325–341. doi: 10.1037/xge0000498
- Hennies, N., Lambon Ralph, M. A., Kempkes, M., Cousins, J. N., and Lewis, P. A. (2016). Sleep spindle density predicts the effect of prior knowledge on memory consolidation. *J. Neurosci.* 36, 3799–3810. doi: 10.1523/JNEUROSCI.3162-15.2016
- Hirshman, E., Trembath, D., and Mulligan, N. (1994). Theoretical implications of the mnemonic benefits of perceptual interference. *J. Exp. Psychol.* 20, 608–620. doi: 10.1037/0278-7393.20.3.608
- Iran-Nejad, A., and Winsler, A. (2000). Bartlett's schema theory and modern accounts of learning and remembering. *J. Mind Behav.* 23, 5–36.
- Jenkins, R., Lavie, N., and Driver, J. (2005). Recognition memory for distractor faces depends on attentional load at exposure. *Psychon. Bull. Rev.* 12, 314–320. doi: 10.3758/BF03196378
- Jiménez, L., Méndez, C., Agra, O., and Ortiz-Tudela, J. (2020). Increasing control improves further control, but it does not enhance memory for the targets in a face-word Stroop task. *Mem. Cogn.* 48, 994–1006. doi: 10.3758/s13421-020-01028-2
- Johnson, J., and Zatorre, R. J. (2006). Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. *Neuroimage* 31, 1673–1681. doi: 10.1016/j.neuroimage.2006.02.026
- Johnson, J. A., Strafella, A. P., and Zatorre, R. J. (2007). The role of the dorsolateral prefrontal cortex in bimodal divided attention: Two transcranial magnetic stimulation studies. *J. Cogn. Neurosci.* 19, 907–920. doi: 10.1162/jocn.2007.19.6.907
- Kalanthroff, E., and Henik, A. (2014). Preparation time modulates pro-active control and enhances task conflict in task switching. *Psychol. Res.* 78, 276–288. doi: 10.1007/s00426-013-0495-7
- Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., et al. (2010). Control and interference in task switching—A review. *Psychol. Bull.* 136, 849–874. doi: 10.1037/a0019842
- Krebs, R. M., Boehler, C. N., De Belder, M., and Egner, T. (2015). Neural conflict-control mechanisms improve memory for target stimuli. *Cerebr. Cortex* 25, 833–843. doi: 10.1093/cercor/bht283
- LaPointe, M. R. P., Rosner, T., Ortiz-tudela, J., Lorentz, L., and Milliken, B. (2022). The attentional boost effect and perceptual degradation : Assessing the influence of attention on recognition memory. *Front. Psychol.* 13, 1024498. doi: 10.3389/fpsyg.2022.1024498
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *J. Exp. Psychol.* 21, 451–468. doi: 10.1037/0096-1523.21.3.451
- Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends Cogn. Sci.* 9, 75–82. doi: 10.1016/j.tics.2004.12.004
- Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Curr. Direct. Psychol. Sci.* 19, 143–148. doi: 10.1177/0963721410370295
- Lavie, N., and Cox, S. (1997). On the efficiency of visual selective attention: Efficient visual search leads to inefficient distractor rejection. *Psychol. Sci.* 8, 395–396. doi: 10.1111/j.1467-9280.1997.tb00432.x
- Lavie, N., Lin, Z., Zokaei, N., and Thoma, V. (2009). The role of perceptual load in object recognition. *J. Exp. Psychol.* 35, 1346. doi: 10.1037/a0016454
- Maddox, G. B., and Balota, D. A. (2015). Retrieval practice and spacing effects in young and older adults: An examination of the benefits of desirable difficulty. *Mem. Cogn.* 43, 760–774. doi: 10.3758/s13421-014-0499-6
- McClelland, J. L. (2013). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *J. Exp. Psychol.* 142, 1190–1210. doi: 10.1037/a0033812
- Meier, B., and Muhmenthaler, M. C. (2021). Different impact of perceptual fluency and schema congruency on sustainable learning. *Sustainability* 13, 7040. doi: 10.3390/su13137040
- Middlebrooks, C. D., Kerr, T., and Castel, A. D. (2017). Selectively distracted: Divided attention and memory for important information. *Psychol. Sci.* 28, 1103–1115. doi: 10.1177/0956797617702502
- Monsell, S., and Driver, J. (2000). *Control of Cognitive Processes: Attention and Performance XVIII*, Vol. 18. MIT Press.
- Moscovitch, M., Rosenbaum, R. S., Gilboa, A., Addis, D. R., Westmacott, R., Grady, C., et al. (2005). Functional neuroanatomy of remote episodic, semantic and spatial memory: A unified account based on multiple trace theory. *J. Anat.* 207, 35–66. doi: 10.1111/j.1469-7580.2005.00421.x
- Muhmenthaler, M. C., and Meier, B. (2019). Task switching hurts memory encoding. *Exp. Psychol.* 66, 58–67. doi: 10.1027/1618-3169/a000431
- Muhmenthaler, M. C., and Meier, B. (2021a). Different impact of task switching and response-category conflict on subsequent memory. *Psychol. Res.* 85, 679–696. doi: 10.1007/s00426-019-01274-3
- Muhmenthaler, M. C., and Meier, B. (2021b). Response-category conflict improves target memory in a flanker paradigm. *Memory* 2021, 1–8. doi: 10.1080/09658211.2021.2012580
- Muhmenthaler, M. C., and Meier, B. (2022). Attentional attenuation (rather than attentional boost) through task switching leads to a selective long-term memory decline. *Front. Psychol.* 2022, 6709. doi: 10.3389/fpsyg.2022.1027871
- Mulligan, N. W., Spataro, P., and Picklesimer, M. (2014). The attentional boost effect with verbal materials. *J. Exp. Psychol.* 40, 1049–1063. doi: 10.1037/a0036163
- Naveh-Benjamin, M., Craik, F. I. M., Guez, J., and Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. *J. Exp. Psychol.* 24, 1091–1104. doi: 10.1037/0278-7393.24.5.1091
- Naveh-Benjamin, M., Guez, J., and Marom, M. (2022). The effects of divided attention at encoding on item and associative memory. *Mem. Cogn.* 50, 59–76. doi: 10.3758/bf03196123
- Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., and Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn. Affect. Behav. Neurosci.* 12, 241–268. doi: 10.3758/s13415-011-0083-5
- Nussenbaum, K., Amso, D., and Markant, J. (2017). When increasing distraction helps learning: Distractor number and content interact in their effects on memory. *Attent. Percept. Psychophys.* 79, 2606–2619. doi: 10.3758/s13414-017-1399-1
- Oberauer, K. (2019). Working memory and attention - A conceptual analysis and review. *J. Cogn.* 2, 58. doi: 10.5334/joc.58
- Olivers, C. N. L., and Meeter, M. (2008). A boost and bounce theory of temporal attention. *Psychol. Rev.* 115, 836–863. doi: 10.1037/a0013395
- Ortiz-Tudela, J., Milliken, B., Jiménez, L., and Lupiáñez, J. (2018). Attentional influences on memory formation: A tale of a not-so-simple story. *Mem. Cogn.* 46, 544–557. doi: 10.3758/s13421-017-0784-2
- Ortiz-Tudela, J., Milliken, B., Botta, F., LaPointe, M., and Lupiáñez, J. (2017). A cow on the prairie vs. a cow on the street: Long-term consequences of semantic conflict on episodic encoding. *Psychol. Res.* 81, 1264–1275. doi: 10.1007/s00426-016-0805-y
- Otten, L. J., Henson, R. N. A., and Rugg, M. D. (2002). State-related and item-related neural correlates of successful memory encoding. *Nat. Neurosci.* 5, 1339–1344. doi: 10.1038/nn967
- Poppenk, J., Köhler, S., and Moscovitch, M. (2010). Revisiting the novelty effect: When familiarity, not novelty, enhances memory. *J. Exp. Psychol.* 36, 1321–1330. doi: 10.1037/a0019900

- Ptok, M. J., Hannah, K., and Watter, S. (2021). Memory effects of conflict and cognitive control are processing stage-specific: Evidence from pupillometry. *Psychol. Res.* 85, 1029–1046. doi: 10.1007/s00426-020-01295-3
- Ptok, M. J., Thomson, S. J., Humphreys, K. R., and Watter, S. (2019). Congruency encoding effects on recognition memory are processing stage specific. *Front. Psychol.* 10, 858. doi: 10.3389/fpsyg.2019.00858
- Quent, J. A., Greve, A., and Henson, R. N. (2021). Shape of U: The relationship between object-location memory and expectedness. *Psychol. Sci.* 2021, 1–14. doi: 10.31234/osf.io/xq37j
- Reggev, N., Sharoni, R., and Maril, A. (2018). Distinctiveness benefits novelty (and not familiarity), but only up to a limit: The prior knowledge perspective. *Cogn. Sci.* 42, 103–128. doi: 10.1111/cogs.12498
- Reynolds, J. R., Donaldson, D. I., Wagner, A. D., and Braver, T. S. (2004). Item- and task-level processes in the left inferior prefrontal cortex: Positive and negative correlates of encoding. *NeuroImage* 21, 1472–1483. doi: 10.1016/j.neuroimage.2003.10.033
- Rhodes, M. G., and Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *J. Exp. Psychol. Gen.* 137, 615–25. doi: 10.1037/a0013684
- Richter, F. R., and Yeung, N. (2012). Memory and cognitive control in task switching. *Psychol. Sci.* 23, 1256–1263. doi: 10.1177/0956797612444613
- Richter, F. R., and Yeung, N. (2015). Corresponding influences of top-down control on task switching and long-term memory. *Quart. J. Exp. Psychol.* 68, 1124–1147. doi: 10.1080/17470218.2014.976579
- Roediger, H. L., and Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychol. Sci.* 17, 249–255. doi: 10.1111/j.1467-9280.2006.01693.x
- Rogers, R. D., and Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *J. Exp. Psychol.* 124, 207. doi: 10.1037/0096-3445.124.2.207
- Rohrer, D., and Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instruct. Sci.* 35, 481–498. doi: 10.1007/s11251-007-9015-8
- Rosner, T. M., D'Angelo, M. C., MacLellan, E., and Milliken, B. (2015a). Selective attention and recognition: effects of congruency on episodic learning. *Psychol. Res.* 79, 411–424. doi: 10.1007/s00426-014-0572-6
- Rosner, T. M., Davis, H., and Milliken, B. (2015b). Perceptual blurring and recognition memory: A desirable difficulty effect revealed. *Acta Psychol.* 160, 11–22. doi: 10.1016/j.actpsy.2015.06.006
- Rummer, R., Schweppe, J., and Schwede, A. (2016). Fortune is fickle: Null-effects of disfluency on learning outcomes. *Metacogn. Learn.* 11, 57–70. doi: 10.1007/s11409-015-9151-5
- Seufert, T., Wagner, F., and Westphal, J. (2017). The effects of different levels of disfluency on learning outcomes and cognitive load. *Instruct. Sci.* 45, 221–238. doi: 10.1007/s11251-016-9387-8
- Slamecka, N. J., and Graf, P. (1978). The generation effect: Delineation of a phenomenon. *J. Exper. psychol. Human. Learn. Memory.* 4, 592.
- Staresina, B. P., and Davachi, L. (2006). Differential encoding mechanisms for subsequent associative recognition and free recall. *J. Neurosci.* 26, 9162–9172. doi: 10.1523/JNEUROSCI.2877-06.2006
- Staresina, B. P., Gray, J. C., and Davachi, L. (2009). Event congruency enhances episodic memory encoding through semantic elaboration and relational binding. *Cerebr. Cortex* 19, 1198–1207. doi: 10.1093/cercor/bhn165
- Sungkhasettee, V. W., Friedman, M. C., and Castel, A. D. (2011). Memory and metamemory for inverted words: Illusions of competency and desirable difficulties. *Psychon. Bullet. Rev.* 18, 973. doi: 10.3758/s13423-011-0114-9
- Swallow, K. M., and Jiang, Y. V. (2010). The attentional boost effect: Transient increases in attention to one task enhance performance in a second task. *Cognition* 115, 118. doi: 10.1016/j.cognition.2009.12.003
- Swallow, K. M., and Jiang, Y. V. (2013). Attentional load and attentional boost: A review of data and theory. *Front. Psychol.* 4, 274. doi: 10.3389/fpsyg.2013.00274
- Tsal, Y., and Benoni, H. (2010). Diluting the burden of load: Perceptual load effects are simply dilution effects. *J. Exp. Psychol.* 36, 1645–1656. doi: 10.1037/a0018172
- Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., et al. (2007). Schemas and memory consolidation. *Science* 316, 1135935. doi: 10.1126/science.1135935
- Tulving, E. (1985). Memory and consciousness. *Can. Psychol.* 26, 1. doi: 10.1037/h0080017
- Uncapher, M. R., and Rugg, M. D. (2005). Effects of divided attention on fMRI correlates of memory encoding. *J. Cogn. Neurosci.* 17, 1923–1935. doi: 10.1162/089892905775008616
- Van Kesteren, M., Beul, S., Takashima, A., Henson, R., Ruiter, D. J., and Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. *Neuropsychologia* 51, 2352–2359. doi: 10.1016/j.neuropsychologia.2013.05.027
- Van Kesteren, M., and Meeter, M. (2020). How to optimize knowledge construction in the brain. *Sci. Learn.* 20, 64. doi: 10.1038/s41539-020-0064-y
- Van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., and Fernández, G. (2013). Consolidation differentially modulates schema effects on memory for items and associations. *PLoS ONE* 8, e56155. doi: 10.1371/journal.pone.0056155
- Van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., and Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends Neurosci.* 35, 211–219. doi: 10.1016/j.tins.2012.02.001
- von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld [The effects of field formation in the trace field]. *Psychologie Forschung.* 18, 299–342.
- Walczyk, J. J. (2000). The interplay between automatic and control processes in reading. *Read. Res. Quart.* 35, 554–566. doi: 10.1598/RRQ.35.4.7
- Wang, S.-H., and Morris, R. G. M. (2010). Hippocampal-neocortical interactions in memory formation, consolidation, and reconsolidation. *Ann. Rev. Psychol.* 61, 49–79. doi: 10.1146/annurev.psych.093008.100523
- Weissgerber, S. C., and Reinhard, M.-A. (2017). Is disfluency desirable for learning? *Learn. Instr.* 49, 199–217. doi: 10.1016/j.learninstruc.2017.02.004
- Woodward, T. S., Meier, B., Tipper, C., and Graf, P. (2003). Bivalency is costly: Bivalent stimuli elicit cautious responding. *Exper. Psychol.* 50, 233–238. doi: 10.1026//1618-3169.50.4.233
- Wylie, G., and Allport, A. (2000). Task switching and the measurement of “switch costs”. *Psychol. Res.* 63, 212–233. doi: 10.1007/s004269900003
- Yeung, N., Botvinick, M. M., and Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychol. Rev.* 111, 931. doi: 10.1037/0033-295X.111.4.931
- Yeung, N., Nystrom, L. E., Aronson, J. A., and Cohen, J. D. (2006). Between-task competition and cognitive control in task switching. *J. Neurosci.* 26, 1429–1438. doi: 10.1523/JNEUROSCI.3109-05.2006
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *J. Mem. Lang.* 46, 441–517. doi: 10.1006/jmla.2002.2864
- Yue, C. L., Castel, A. D., and Bjork, R. A. (2013). When disfluency is—and is not—a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Mem. Cogn.* 41, 229–241. doi: 10.3758/s13421-012-0255-8