A functional perspective on schema-based learning and recognition of novel word associations

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
der Fakultät HW
Bereich Empirische Humanwissenschaften
der Universität des Saarlandes

vorgelegt von

Julia Anuschka Meßmer

aus Dudweiler

Saarbrücken, 2024

Dekan:
Prof. Dr. Peter Loos, Universität des Saarlandes

Berichterstatter:
Prof. Dr. Axel Mecklinger, Universität des Saarlandes
Prof. Dr. Dirk Wentura, Universität des Saarlandes

Tag der Disputation:
20.07.2023

Zusammenfassung

Das Ziel der vorliegenden Arbeit war es, eine funktionelle Perspektive auf das schema-basierte Lernen neuer Wortassoziationen (Komposita) und deren späteres Wiedererkennen zu entwickeln. Dazu wurden zwei Forschungsideen zusammengeführt. Da sowohl schema-basiertes Lernen (z.B., Hebscher et al., 2019; van Kesteren et al., 2012) als auch Unitarisierung (z.B., Bader et al., 2014; Haskins et al., 2008; siehe auch Henke, 2010) weniger hippocampale Beteiligung aufweisen als traditionelles Assoziationslernen, formulierten wir die Hypothese, dass Schemakongruenz die Bildung unitarisierter Repräsentationen unterstützen könnte, die dann mittels eines absoluten Vertrautheitsprozesses wiedererkannt werden könnten (Mecklinger & Bader, 2020).

Die drei Experimente, die in der vorliegenden Arbeit dargestellt sind, beinhalten alle eine inzidentelle Lernphase, in der neue Komposita zusammen mit einer kongruenten oder neutralen vorangehenden Definition gelernt wurden (experimentelle Manipulation von Schemakongruenz). Nach einem Retentionsintervall von etwa 10 Minuten folgte ein überraschender, nicht vorangekündigter Gedächtnistest. In dieser Testphase sahen die Teilnehmenden verschiedene Arten von Komposita und sollten diese als intakt, rekombiniert oder neu klassifizieren (Experiment 1), als alt (intakt) oder neu (rekombiniert, ähnliche Distraktoren; Experiment 3) oder bearbeiteten eine lexikalische Entscheidungsaufgabe (Experiment 2).

Unsere Ergebnisse implizieren, dass drei Prozesse am schema-basiertem Lernen beteiligt sind. Semantisches Priming, angezeigt durch eine reduzierte N400 Amplitude in der schema-kongruenten Bedingungen, führt zu Schemakongruenz. Die bedingungsunabhängige semantische Integration der Wortbestandteile ist förderlich für die Gedächtnisbildung, indiziert durch einen N400 *Subsequent Memory Effect* (SME). Der dritte Prozess, die schemakongruenzgetriebene Bildung einer konzeptuellen (unitarisierten) Repräsentation wird angezeigt durch einen größeren parietalen SME in der kongruenten im Vergleich zur neutralen Bedingung.

Basierend auf dem behavioralen Ergebnismuster, dass assoziatives Gedächtnis mehr von Schemakongruenz profitiert als Itemgedächtnis (siehe auch Parks & Yonelinas, 2015), könnte Schemakongruenz die Bildung von

unitarisierten Repräsentationen fördern. Die neurokognitiven Prozesse, die dem Wiedererkennen solcher Komposita unterliegen, beinhalten wahrscheinlich einen höheren Anteil absoluter Vertrautheit in der kongruenten als in der neutralen Bedingung, indiziert durch einen entsprechenden reduzierten N400-Effekt. Basierend auf den Ergebnissen des dritten Experiments, bei dem der Rekognitionstest semantisch ähnliche Distraktoren beinhaltete, schlussfolgerten wir, dass die Repräsentationen, die unter dem Einfluss eines Schemas gebildet werden, detailarm sind und lediglich die semantische Konzeptstruktur (*gist*) beinhalten. Diese Repräsentationen könnten parallel zu episodischen Assoziationen geformt werden, die wahrscheinlich beim traditionellen Assoziationslernen gebildet werden. Die unitarisierten Repräsentationen konnten hierbei nicht nur in einem expliziten Gedächtnistest verwendet werden, sondern auch die Performanz in einer impliziten Gedächtnisaufgabe beeinflussen.

Abstract

With the current research, we sought to develop a functional perspective on schema-based learning of novel word associations, i.e., novel compound words and their later recognition. In combining the idea that both, schema-based learning (e.g., Hebscher et al., 2019; van Kesteren et al., 2012) and unitization (e.g., Bader et al., 2014; Haskins et al., 2008; see Henke, 2010) might rely less on hippocampal contribution than traditional associative learning, we hypothesized that schema-congruency might support the formation of unitized representations that could then be recognized by means of an absolute familiarity process (Mecklinger & Bader, 2020).

All three experiments presented include an incidental learning phase, in which novel compound words were learned together with a preceding definition that was either congruent or neutral (experimental manipulation of schema congruency). After a retention interval of about 10 minutes, a surprise memory test followed. In the test phase, participants were shown different types of compound words and instructed to classify each as intact, recombined, or new (Exp. 1), as old (intact) or new (recombined, similar lures; Exp. 3) or underwent an implicit lexical decision task (Exp. 2).

Our results imply that three processes might underly schema-based learning. Semantic priming, indicated by an N400 attenuation effect in the schema-congruent condition, establishes schema congruency. Condition-independent semantic integration of the constituents is beneficial for memory formation, as indicated by an N400 subsequent memory effect (SME). Lastly, we found a larger parietal SME in the congruent than in the neutral condition. This might reflect the formation of a conceptual (unitized) representation under the influence of a congruent schema.

Second, based on our results, schema-congruency might support the formation of unitized representations, indicated by schema-congruency being more beneficial for associative than item memory performance (see Parks & Yonelinas, 2015). The neurocognitive processes underling recognition of those compound words might include larger absolute familiarity contributing to associative recognition in the congruent than in a neutral control condition, indicated by an N400 attenuation effect. Based on data from our third experiment

including semantically similar distractors during the recognition memory test, we concluded that the representations formed under the influence of a schema might be gist-like. Those might be created next to episodic associations that are probably also formed in traditional associative learning. Lastly, those unitized memory representations formed under the influence of a schema cannot only be accessed in an explicit memory test, but also affect performance in an implicit memory test.

Acknowledgements

Although the current thesis is written by my own, this product would not have been possible without the valuable help of many people. I experienced at first hand that teamwork is an integral part of science, and I am forever grateful to all the people who have accompanied me on this journey.

First and foremost, I would like to thank my supervisor Professor Dr. Axel Mecklinger for the opportunity to do research in such a fascinating field, for always being supportive and available for discussions, for letting me pursue own scientific ideas and for an always open door. I am also grateful to Dr. Regine Bader for her valuable contribution to the research presented here, for many fruitful discussions and her help in technical questions. Sometimes, especially at the beginning, science can feel like a labyrinth. I am grateful for Professor Dr. Axel Mecklinger and Dr. Regine Bader helping me to find my way through. I am also grateful to Professor Dr. Dirk Wentura for reviewing this thesis.

I also would like to thank my former and current colleagues at the Experimental Neuropsychology Unit, for their helpful feedback, for many fruitful discussions, for their support and lunch/coffee-breaks that I highly appreciated. Thank you for the many compound words you invented throughout this journey ;-) (that unfortunately did not make it into the experiments)! I am also grateful for the opportunity to pursue a PhD in such a great working environment as it is provided by the SFB1102. I am pretty sure that there are close-to-zero ideas presented in this thesis that have not benefitted from discussion with at least one of my colleagues (from the SFB 1102 or the Experimental Neuropsychology Unit). I would also like to thank the students, student research assistants, and interns who contributed to this project. Thank you to all volunteers who participated in the experiments presented here. I also would like to thank the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft[Project-ID 232722074 – SFB 1102, Project A6]) for funding this research as part of the SFB 1102.

Being now nearly at the end of this journey, I must conclude that some ways are a lot easier and have a much more beautiful view with travel buddies.

I would like to thank Dr. Véronique Huffer, the second constituent of the Kreativbüro-Unit, for her enormous support, the many discussions, the creative breaks, the fun and also for her helpful comments on parts of this thesis.

I would also like to thank Linda Sommerfeld. We have been study buddies since undergrads and being PhD buddies seems to be an appropriate final for what started as a learning group. Thank you for making this journey more enjoyable and a lot easier.

I am deeply grateful to my friends and family and especially to my parents Anne Meßmer and Stephan Meßmer for their unconditional support in life and this thesis. Lastly, I am grateful to Max for helping me through stressful times.

Peer-reviewed articles and manuscripts

This dissertation is built on two original empirical articles of which one is published in an international, peer-reviewed journal, whereas the other is submitted. Whilst I am the first author of both, other people significantly contributed to this work as co-authors and their contribution is indicated below. A slightly modified version of the first article forms chapter 4 of the current dissertation, whereby single paragraphs have been (re)moved to chapters 1, 2 and 3 for reasons of text coherence. A similar procedure was applied for the second article which underlies chapter 5 and from which also single paragraphs have been used in chapters 1, 2 and 3.

Article 1

Meßmer, J. A., Bader, R., & Mecklinger, A. (2021). The more you know: Schema-congruency supports associative encoding of novel compound words. Evidence from event-related potentials. Brain and Cognition, 155, 105813. https://doi.org/10.1016/j.bandc.2021.105813

Article 2

Meßmer, J. A., Bader, R., & Mecklinger, A. The more you now: Schema-congruency supports the formation of unitized representations. Evidence from event-related potentials. [Manuscript submitted for publication]. Experimental Neuropsychology Unit, Saarland University

As the contribution of the co-authors to the chapters is outlined here, I use the more convenient “we” form throughout the manuscript.

Table of Contents

Figures

Tables

Abbreviations

ANOVAs, analyses of variances

CARIN, Competition Among Relations in Nominals (theory)

CLST, complementary learning systems theory

DRM, Deese-Roediger-McDermott (task)

EEG, electroencephalography

ERPs, event-related potentials

LOU, levels-of-unitization

LPC, late parietal component

MANOVAs, multivariate analyses of variance

mPFC, medial prefrontal cortex

MTL, medial temporal lobe

PA, paired associate

RI, retrieval integration (account)

ROCs, receiver operating characteristics

RQs, research question

SLIMM, schema-linked interactions between medial prefrontal and medial temporal regions

SME, subsequent memory effect

vmPFC, ventromedial prefrontal cortex

0 Introduction

Do you know what a snowy owl is? You probably do. However, do you remember when and where you initially learned what a snowy owl is? This one might be harder to remember, as you probably stumble over this (compound) word so many times that you lost the memory of your first encounter with this word. But how is it even possible that we remember concepts acquired over many years and have the capability to learn new ones, as for example the novel word *Starchair*, reflecting a new concept?

This is subject to the assumption that we bear a permanent storage, „a system or systems assumed to underpin the capacity to store information over longer periods of time “(Baddeley, 2020, p. 13) of what we experienced with us, so-called long-term memory. Our memory system is thought to contain a collection of memory representations[1] of past episodes (e.g., the wedding of a friend) and structured knowledge (e.g., a snowy owl has white feathers with black spots on it), which we can (with certain limitations) access as desired. After information has been encoded into long-term memory, laying down a memory representation, a stabilizing consolidation process is thought to occur (Dudai, 2012; Frankland & Bontempi, 2005; Squire & Alvarez, 1995) and the memory can later be retrieved. But how is long-term memory organized and how can we use what we already know (what is a chair, what is a star) to learn the novel word *Starchair*? Learning Starchair requires learning a novel association between the already known words *Star* and *Chair*.

One approach to understand a scientific concept is to analyze its components (see Titchener, 1898). Within such structuralist, systems-oriented approaches to the psychology of memory, it has been proposed that long-term memory constitutes subordinate memory systems, differing in their content and also in functional properties (see Squire & Zola, 1996, for a taxonomy). In line with this, representations of perceived episodes (e.g., having read Starchair in the newspaper) are assigned to episodic memory, which "receives and stores information about temporally dated episodes or events, and temporal-spatial

[1] Representationalism is a philosophical view "according to which remembering is mediated by representations whose contents represent past intentional objects" (De Brigard, 2014, p. 402).

relations among these events" (Tulving, 1972, p. 385). Episodic memory has a special phenomenological quality, as we feel we can mentally travel back in time and relive those memories in our mind (Tulving, 1983; 1985; 2002).

Critically, episodic memories are neither relived nor acquired in isolation. Our memory system is not only a rather loose and bulky collection of single memories in our head, but consists of structured knowledge, linking single memories, which we can flexibly access at will (see Bartlett, 1932). In the systems-oriented approach, this extracted prior knowledge has been referred to as semantic memory, "the memory necessary for the use of language. It is a mental thesaurus, organized knowledge a person possesses about words and other verbal symbols, their meaning and referents, about relations among them, and about rules, formulas, and algorithms for the manipulation of these symbols, concepts, and relations" (Tulving, 1972, p. 386). Knowing what Stars and Chairs are, for example, is part of semantic memory. Similar to episodic memory, semantic memory is built from associations that some argued form some kind of network structure with nodes for concepts and associations between nodes for interrelations (Collins & Loftus, 1975).

Research converged on the idea that storing episodic memories relies on a particular brain structure within the medial temporal lobe, the hippocampus, which is capable to rapid learning (e.g., Eichenbaum et al. 2007; McClelland et al., 1995). In contrast, it has been proposed that acquiring semantic associations takes time, as it requires either many learning instances, i.e., many repetitions (McClelland et al., 1995), or occurs via consolidation of episodic (hippocampal) associations (Dudai, 2012; Frankland & Bontempi, 2005; McClelland et al., 1995; Squire & Alvarez, 1995). Thus, whilst initially the information that Starchair was read in a newspaper is available (episodic), it is lost over time and only the word form together with its meaning is restored in semantic memory.

The latter point shows that episodic and semantic memory interact whenever information is processed or reconstructed (see Barry & Maguire, 2019), which is then interpreted and stored in relation to our existing semantic knowledge (Gilboa & Marlatte, 2017). In the Starchair example, when the new word is read and processed for the first time, the reader probably activates the concepts *Star* and *Chair* in order to make sense of it. Second, and in line with this, it is assumed that episodic memories are stored within an ensemble, held together with knowledge structures (Bartlett, 1932; McClelland et al., 1995). This means that after learning, the novel word Starchair should be embedded into our semantic knowledge and connected to the concepts *Star* and *Chair*. But how exactly are those connections between existing knowledge and incoming information formed?

This question has been addressed within schema theories, having a long history in cognitive psychology, dating back to Bartlett (1932) and Piaget (1926). Progress in cognitive neuroscience and the development of theories led to their reinvention in neuroscience (Tse et al., 2007; 2011). This resurrecting interest probably originates from the idea that schema-based learning provides an exception from the rule that rapid learning of associations relies on the hippocampus (see Hebscher et al., 2019; van Kesteren et al., 2012), which might indicate a qualitatively different learning mechanism underlying schema-based learning.

Motivated by the insights from this rather structuralist brain system-oriented approach, the goal of the current thesis was to shed light on prior-knowledge based learning of novel associations from a functional, cognitive process level. Hereby, we sought to unravel the neurocognitive processes involved in schema-based learning of completely new associations that bear a novel concept, i.e., novel compound words as *Starchair*, and their later recognition (see Mandler, 1980). In research on associative recognition, two processes, familiarity and recollection have been identified (Mandler, 1980; Yonelinas, 2002; 2010). Whilst it is assumed that (hippocampal) recollection is typically required to recognize associations, familiarity can underly associative recognition when unitized representations of the association have been formed (e.g., Bader et al., 2010; 2014; Wiegand, 2010; see Yonelinas, 2002). Based on a process model on the learning of associations (Henke, 2010), in the current thesis, we investigated whether the formation of unitized representations might be a possible mechanism underlying schema-based learning of novel associations.

To make inferences about cognitive processes, we used the experimental method and event-related potentials (ERPs). ERPs are changes in voltage that are generated by the brain (Hagoort, 2003; Luck, 2014), measured on the surface of the scalp and time-locked to an event. Hereby, the events can be external, i.e., evoked by the environment, or internal, generated by cognitive operations (Hagoort, 2003; Luck, 2014; Payne et al., 2020). ERPs are an interesting tool for the investigation of cognitive processes as these electrical current changes propagate instantaneously to the scalp, leading to a millisecond-level temporal resolution (Payne et al., 2020).

In this thesis, we addressed the following three research questions concerning how congruency with a schema, influences associative memory. First, what are the neurocognitive mechanisms driving schema-based learning of novel compound words? Second, what type of representation is formed and what neurocognitive processes underly recognition? Lastly, how can these representations be retrieved? We addressed those research questions in presenting empirical data from three experiments.

1 The many facets of the schema concept and its role for memory research

New information is not acquired in isolation but rather influenced by the organized mass of knowledge extracted from our accumulated experiences (Bartlett, 1932). These knowledge structures (schemata) are not only referred to when new information is acquired, but also permanently in use to interpret the unfolding reality and react on it by selecting appropriate actions (Gilboa & Marlatte, 2017). Common operationalizations of schema knowledge are for example concept or category knowledge (Höltje et al., 2019; Höltje & Mecklinger, 2022; Schulman, 1974), abstract rules about the value of every-day objects (Greve et al., 2019), conceptual knowledge contained in scene-object relationships (van Kesteren et al., 2013) or in a movie (van Kesteren et al., 2010). How is learning of novel associations achieved by means of these knowledge structures?

1.1 The schema concept through the ages

The *schema concept* has a long history in philosophical and psychological research, although its meaning varies extensively depending on the theories within which it is used. Except for a few earlier mentions (e.g., Head & Holmes, 1911; Piaget, 1926), the schema concept gained popularity in cognitive research being introduced within Bartlett's (1932) theory of remembering. In his functional approach, Bartlett defines a schema as a transiently active functional pattern (Bartlett, 1932; see Iran-Nejad, 1980; Iran-Nejad & Winsler, 2000). He proposes his theory as an opponent view to what Wagoner (2013) calls the trace theory of memory. The trace theory is very popular until today and rests on the assumption that memories are in any form stored in the head (e.g., as a memory trace, representation or engram; Bartlett, 1932; see Wagoner, 2013). Bartlett (1932) highlights the dynamic nature of schemas (see Iran-Nejad & Winsler, 2000) and defines it as "an active organisation of past reactions, or of experiences" (p. 210), that should operate in well-adapted organic responses. Thereby, a schema renders a specific reaction possible with a preceding incoming impulse.

After Bartlett's theory gained popularity, the epistemological character of the schema concept changed towards a rather structural approach in line with the trace assumption (see Iran-Nejad, 1980; Iran-Nejad & Winsler, 2000; Wagoner, 2013). This approach informed many cognitive theories in the 1970s and 1980s dedicated to questions about how comprehension and learning work, taking into account prior knowledge (e.g., Alba & Hasher, 1983; Rumelhart, 1980; Rumelhart & Ortony, 1977). Iran-Nejad & Winsler (2000) state that those type of structuralist schema theories (e.g., Rumelhart, 1980; Rumelhart & Ortony, 1978) see "knowledge construction as the retrieval and instantiation of pre-existing long-term memory schemas" (i.e., prior knowledge; p. 15). They argue that the "most severe problem facing the structural schema approach concerns learning" (p. 15) especially how schema-inconsistent information is learned (see also van Kesteren et al., 2012, Box 1; but see Graesser et al., 1979; Schank, 1982; for schema models including learning of inconsistent information).

The schema concept is woven through the literature of several different fields as e.g., cognitive psychology (Neisser, 1967), artificial intelligence (Rumelhart, 1980) and education (McVee et al., 2005). The heterogeneity in use of the schema concept has led to the criticism that it is ill-defined (e.g., Bartlett, 1932; Gosh & Gilboa, 2014) and disintegrated over years, lacking a unified account (van Kesteren, 2013). Now, almost ten years later, a unified account of schema-based learning is still waiting to be developed. However, an important argument brought forward by Fiske & Linville (1980) is that the schema concept alone has little if any explanative value, but rather should be seen as a meta-construct, the explanative power of which depends on the general cognitive framework in which it is put in. In the current work, we hold the view that the newer neuroscientific approach to the schema concept, embedded in the current popular neurocognitive memory framework, has the potential to provide a basis for understanding how prior knowledge interacts with the learning of new associations. Including neuroscientific research is crucial, as cognition is closely tied to the brain and thus, a unifying account of schema theory would require a plausibility check from neuroscientific progress (van Kesteren, 2013). In line with this, the focus of the schema concept has shifted towards a neuronal, brain-systems oriented approach with special interest on the question how prior knowledge informs learning (Gilboa & Marlatte, 2017; Gosh & Gilboa, 2014; van Kesteren et al., 2012).

1.2 The general neurocognitive framework in which the schema concept is embedded

Today’s cognitive neuroscience theories are mostly derivates of the memory trace idea, meaning that memories or related structures are to some extent stored in the brain which can be used to re-activate or reconstruct the memory (see De Brigard, 2014; Wagoner, 2013). This has been referred to as memory trace, memory representation or engram with the idea that there is a neurochemical substrate of memories in the brain (Lashley, 1980; Semon, 1904; see Hebscher et al., 2019). Here, an initial trigger or cue is thought to initiate a sequence of cellular and molecular changes within and across brain cells. Those give rise to the creation of a distributed trace or engram which temporally outlasts the triggering event (Takeuchi et al., 2022). Memory is thought to be constructive (Barry & Maguire, 2019; Bartlett, 1932), meaning that a memory representation does not contain an exact image of reality as it is but rather holds an image of reality as we perceive it – our interpretations of reality (Neisser, 1967, p. 302).

1.2.1 Episodic or Semantic?

As already elaborated on in the introduction, memories and their underlying traces or representations can be classified as episodic or semantic (Tulving, 1972). Hereby, slightly different connotations of the labels are in use. In the strict sense, episodic memories can be delineated from semantic memories by the availability of (spatiotemporal) context information (Eichenbaum et al., 2007; Preston & Eichenbaum, 2013; see Tulving, 1972). In a wider sense, learning in the laboratory setting is often ascribed to episodic memory as the memory test asks for an information bound to an experimental episode (Anderson, 2020, p. 170). However, this answer only targets one aspect of characterizing episodic and semantic memory and it is unclear whether all memory traces formed within a laboratory experiment necessarily contain a viable link to the spatiotemporal context of the experiment, justifying their episodic character, or are semantic in nature (see Tulving, 2002 for a similar argument). If the memory test does not directly assess the context information, this question cannot be answered by just referring on this logic and the question whether episodic or semantic memory has been acquired in an experiment is rather an empirical than a theoretical question. Thus, in the current thesis, the labels episodic and semantic are used to refer to properties of the memory representation, which is the availability of spatiotemporal context information.

In Tulving's (1972) definition of semantic memory, he deems it being closely tied to language. Memory storage of words is thought to be achieved by networks comprising a conceptual level, a lemma level at which grammatical properties are available and a form level, at which morphological information is available (Bock & Levelt, 1994). Thus, word form and semantics are considered to be connected (Poeppel & Idsardi, 2022) and in memory research, both are typically subsumed under semantic memory. For the sake of the current considerations, it is not necessary to increase the resolution of semantic memory and prior knowledge in differentiating between lexical and semantic information. However, those may be interesting topics to address in future research.

1.2.2 Brain systems underlying episodic versus semantic memories

Episodic and semantic memory have also been specified on a brain systems level. Traditionally, memory research converged on the idea that learning of associations relies on the hippocampus (e.g., Davachi et al., 2003; see Henke, 2010), a brain structure located in the medial temporal lobe, in which the respective memory representations are initially stored. Episodic memories including spatiotemporal context information are thought to rely on hippocampal representation (e.g., Eichenbaum et al., 2007; Preston & Eichenbaum, 2013). Hereby, the hippocampus has an ideal anatomical position to bind different types of information: It receives input from its surrounding parahippocampal and perirhinal cortices, that are thought to be responsible for memory of object and scene information, respectively, and also indirect input from different association cortices (see Eichenbaum et al., 2007; Preston & Eichenbaum, 2013; Ranganath & Ritchey, 2012). The neocortex is considered to be a more permanent storage (McClelland et al., 1995), holding semantic memory. The medial temporal lobe system including the hippocampus and the neocortex are thought to interact, whereby the medial temporal lobe is thought to store information pointing to and activating relevant neocortical sites rather than the whole memory representation (Squire & Alvarez, 1995; Squire et al., 1983; Teyler & DiScenna, 1986).

Based on neuropsychological evidence and computational modeling, McClelland and colleagues (1995) propose the existence of two complementary learning systems in the brain with different computational properties: one slow learning system located in neocortical structures and one fast learning system in hippocampal structures. When new, arbitrary elements are added to the system (i.e., inconsistent information), it is assumed that one repetition is not sufficient for the slow learning system to acquire the new information and embed it into the network structure. Rapid learning of arbitrary associations with the computational properties of the neocortical slow learning system would lead to catastrophic interference (McCloskey & Cohen, 1989), meaning that concurring memory traces eliminate each other. The rapid learning system, located in the hippocampus, also includes neural connections that are adjusted during learning, but is capable of rapid learning due to its computational properties. As the systems are connected, hippocampal reinstatement gives also rise to neocortical reinstatement enabling both, task performance and consolidation.

1.2.3 Three stages of a memory representation

Memory formation includes two types of information processing that must be performed by the brain. Those are assigned to two stages, encoding and consolidation. During encoding, an initial memory trace must be formed rapidly. In the second stage, the memory trace is thought to be strengthened, being more stable and durable thereafter (Anderson, 2020, p. 193), and more immune to disruption (Frankland & Bontempi, 2005), an idea which goes back to Müller and Pilzecker, 1900 (see Lechner et al., 1999). Consolidation can complete rather fast on the synaptic level (within hours) and rather slow (within days, weeks, years) on the level of neurobiological systems (see Frankland & Bontempi, 2005; Squire & Alvarez, 1995; Rudy & Sutherland, 2008 for literature reviews; but see

Runyan et al., 2019 for a unified theory of cellular and systems consolidation). In systems consolidation theory, it is assumed that an encoded memory trace initially relies on the hippocampus but is eventually reorganized so that it gradually becomes hippocampus-independent and is assimilated into neocortical structures. Thereafter, cortico-cortical connections can stand on their own. In traditional systems consolidation theory, the psychological nature of the memory traces is thought to remain the same, i.e., the phenomenological quality of remembering which is supported by the memory trace during retrieval. However, alternative models have been proposed in which the hippocampus remains involved (see Gilboa & Moscovitch, 2021; Nadel and Moscovitch, 1997; Winocur et al., 2010; Winocur and Moscovitch, 2011).

The third stage is accessing the content of the memory trace, i.e., retrieval. Episodic memory retrieval is thought to require a cue to reinstate the respective memory trace accompanied by autonoetic awareness, being perceived as mentally traveling back in time (Tulving, 1983). Semantic memory retrieval also requires a cue but is accompanied by noetic awareness (no experience of mental time traveling).

1.3 The neurocognitive schema approach

Within the neuroscientific approach, memory schemas can be defined as “higher-level knowledge structures that organize lower-level representations from long-term memory” (Gilboa & Marlatte, 2017, p. 618). Similar to its predecessors in cognitive science, the schema concept has been used rather loosely in neuroscience and refers to “mental and neurobiological prior associative networks that influence new information processing” (Gilboa & Marlatte, 2017, p. 618). Gilboa and Marlatte compare schemas with “reference templates against which new information can be compared” (p. 618) whereby recurring features can be bound together. Similar to former schema approaches (e.g., Rumelhart, 1980; Rumelhart & Ortony, 1977), they formulate four features of schemata which are that elements bound within a schema are considered nonspecific, lacking unit detail (1) as they reflect commonalities among experiences (2), overlapping and interconnected in an associative network structure (3) and adaptable (4; Ghosh & Gilboa, 2014; Gilboa & Marlatte, 2017;

van Kesteren et al., 2012). In line with Bartletts (1932) schema definition, schemas are constantly evolving dynamic structures, whereby new information is added via an assimilation process and the schema structure is modified via accommodation (Ghosh & Gilboa, 2014; Gilboa & Marlatte 2017; Piaget, 1926).

Thus, a schema is defined as an abstract knowledge structure. In reviewing schema literature, Ghosh & Gilboa (2014) add a conceptual detail and argue that concepts or categories themselves do not fulfil the requirements of a schema (i.e., nonspecific elements, extracted from commonalities, associative network structure, dynamically evolving). Instead, the schema as an abstract superordinate rule can draw on category or conceptual knowledge (Ghosh & Gilboa, 2014; Gilboa & Marlatte, 2017) thus temporarily binding together subordinate representations to form a superordinate knowledge template (Gilboa & Marlatte, 2017). In our operationalization of schemata, we get back to this detail (see chapter 2).

We argued that in the framework by Gilboa & Marlatte (2017), an important feature distinguishing a schema from a loose collection or mere co-activation of semantic associations is its structural capacity to combine several prior known concepts to create new ones (Meßmer et al., 2021; chapter 4), organized within conceptual hierarchies. Thus, a schema provides a framework that structures the processing of already stored information in relation to not yet acquired information. It remains yet unclear how precisely this structuring is achieved on a neuronal level, but it is assumed that the medial prefrontal cortex (mPFC) has a cardinal role as a knowledge hub binding together underlying representations (Gilboa & Marlatte, 2017).

1.3.1 Schema construction and instantiation

Schemas are thought to be built from extracting commonalities across events during consolidation (Gilboa & Marlatte, 2017). Schlichting and Preston (2015) argue that schema extraction is achieved by memory integration, which means that related memories are encoded with overlapping activity patterns.

The schema is activated (instantiated) via a resonance process sensitive to overlap between the schema and incoming information (van Kesteren et al., 2012) or via a context-sensitive process by means of which irrelevant associations are not activated or even inhibited, which is referred to as schema instantiation (Gilboa & Marlatte, 2017; Preston & Eichenbaum, 2013). A schema might be a transient activation of a part of semantic memory which Gilboa and Marlatte (2017) refer to as “co-activated long-term representations” (p. 620). What makes schemas interesting for neuroscience is their role in learning.

1.3.2 Schema-congruent information is remembered better

Schema knowledge does not only influence how information is processed online, but it also impacts which aspects of an event are encoded and retained in memory and which aspects are later forgotten (i.e., Pichert & Anderson, 1977; see Bartlett, 1932; Gilboa & Marlatte, 2017). It is well established that events which are congruent with a given schema are better retained than schema-incongruent events (Alba & Hasher, 1983; Schulman, 1974; Pichert & Anderson, 1977; see Greve et al., 2019, for an overview). The congruency effect has been reported for a wide range of event types and modalities (e.g., face location associations; Atienza et al., 2011; word pairs; Bein et al., 2015; Bein et al., 2014; objects; Greve et al., 2019; word lists; Hall & Geis, 1980; auditory and visual objects; Naghavi et al., 2011; word-color associations; Staresina et al., 2009; associations between visual and tactile information; van Kesteren et al., 2013), although up to now, only few studies have investigated the influence of a memory schema on the learning of associations (e.g., Atienza et al., 2011; Bein et al., 2014; 2015; Staresina et al., 2009; van Kesteren et al., 2013). Staresina and colleagues (2009), operationalized event congruency as the plausibility judgment given by participants for the semantic match of a word–color combination as for example the word elephant printed on a red screen (incongruent) or the word balloon on a yellow screen (congruent). They found superior item memory and also better memory for associated source details (i.e., color information) for congruent events. However, schema congruency is not always beneficial for memory performance.

1.3.3 Schema-congruent information is remembered worse

Schema congruency is not always beneficial for memory performance. Rather, its influence on memory performance depends on the requirements set by the memory test. Prior research provided evidence that it is harder to distinguish highly similar information from originally learned information (Spalding et al., 2015; Sweegers et al., 2015; Webb et al., 2016). This has been traced back to schema representations being less detailed and more gist-like (Gilboa & Marlatte, 2017). Depending on the perspective one likes to take, this can be referred to as *schema generalization* which denotes the idea that a schema is overused (see Bartlett, 1932) and stretched toward a consistent piece of information, which can lead to false memories.

1.3.4 Rapid consolidation of schema-congruent memories?

In a seminal study, Tse and colleagues (2007) trained rats to acquire spatial schemata of flavor-place associations in an event arena. The event arena constitutes an underlying grid with landmarks, in which six sand-wells were positioned at particular locations. In those sand-wells, food pellets of a specific flavor were hidden. In a later non-rewarded cued recall test, rats were given a flavor cue and it was assessed if they preferentially searched (dug) in the respective learned position. In the first experiment, rats with hippocampal lesions were not able to acquire the spatial schema of the event arena in contrast to animals with a non-hippocampal control (sham) lesion. Thus, an intact hippocampus was necessary for initial schema acquisition. In a second experiment, the authors trained another group of rats similarly to experiment 1. Afterwards, two of the six initially learned flavor-place pair-associates were exchanged for two close (schema-consistent) ones, whereby only one rewarded trial was given for each of those. Memory for these new locations was tested in a non-rewarded cued recall trial 24hrs later. This test showed that rats preferentially dug at the new locations when given the appropriate flavor cue, thus showing that they were able to rapidly acquire schema-consistent information. Another 24hrs later, thus 48hrs after initial learning of the two schema-consistent new pair-associates, rats underwent surgery in which lesions were positioned in the hippocampus for one group of rats, whereas lesions were set in non-hippocampal areas for another (sham) group of rats. After having recovered from surgery, rats performed in a series of non-rewarded cued recall trials. Both the group with the hippocampal lesions and the sham group could remember the not modified locations of the initially acquired spatial schema as well as the two new pair-associates that only have been acquired in one rewarded trial. The authors interpreted their results in that information consistent with a schema underlies rapid systems-consolidation, which means that storage and retrieval of the newly learned pair-associates does not rely on the hippocampus anymore, but probably on neocortical structures. The authors also claim that as rapid neocortical consolidation is possible, encoding of associative memory traces must also be rapid, “perhaps even ‘on-line’ within sensory-perceptual systems” (p. 81). However, the respective memory traces are not yet viable on their own, as hippocampal lesions placed 3hrs after initial learning of schema-consistent information disrupts performance in a later memory test. In a later study, Tse and colleagues (2011) use a modified version of the previous experiment to investigate whether in parallel to initial hippocampal encoding, cortical encoding into a schema occurs when new information relates to prior knowledge. Based on their results they conclude that “assimilation of rapidly acquired new PA [paired-associate] information into existing cortically based mental schemas is associated with cortical encoding of information during hippocampal-dependent learning, and that this simultaneous encoding is essential for long-term memory” (p. 893). Thus, studies on spatial schema-based learning in rodents provide evidence that new information is initially dependent on the hippocampus with the prior knowledge-driven parallel formation of a cortical trace as well as rapid consolidation of new schema-congruent information. Research on humans differs from rodent research both in methodology (which is mostly neuroimaging and patient data in humans) as well as in the amount of their prior knowledge, which is much richer in humans (Alonso et al., 2020).

1.3.5 Is there a specific learning mechanism underlying schema-based learning?

Neuroimaging research on humans (e.g., van Kesteren et al., 2010; 2013) provides evidence for a special role of the mPFC in schema-based learning, which is thought to inhibit the medial temporal lobe system when schemata can be used during encoding (van Kesteren et al., 2012) or to hold context-appropriate schemas used by the medial temporal lobe system (Preston & Eichenbaum, 2013).

In reviewing recent research, Hebscher and colleagues (2019) put forward the idea that new cortical engrams may be rapidly viable if they can be associated with active, well-established cortical engrams. One factor promoting this relationship is relatedness to prior knowledge and this idea of a prior-knowledge-driven, distinct learning mechanism dovetails with recent neuroscientific theories on schema-based learning (e.g., van Kesteren et al., 2012; Gilboa & Marlatte, 2017). Hebscher and colleagues argue that amongst other factors, relatedness to prior knowledge can "promote activation of existing neural representations, which facilitates the formation of new cortical associations" (p. 3). Hereby, multimodal information might be brought online (see McKenzie et al., 2014; Preston & Eichenbaum 2013 for a different view on hippocampal involvement). This idea is also in line with the complementary learning systems theory (CLST; McClelland et al., 1995): McClelland (2013) argues that the original CLST was proposed based on modeling of inconsistent information. Following new computational modeling, McClelland (2013) concludes that consistent information can be assimilated rapidly without suffering from heavy interference.

Supporting this idea, Tompary and colleagues (2020) disentangled the accessibility of a schematic memory (i.e., the schema is formed) from its expression (i.e., the schema can exert an influence episodic memories) in a clever behavioral experiment. They used a spatial learning task, in which the spatial position of images of different categories (animals and objects e.g., sunglasses) on a circle was to be learned. This allowed for acquiring of both episodic memories (exact positions) and schematic memories (likely positions as a continuous variable). There was an immediate and a delayed retrieval test either 24hrs or 1 week later. The results show that as expected, precision of episodic memory declined over time (after one week), whilst the influence of schema consistency on episodic retrieval increased over time (after one week) and was larger for weaker memories (low confidence). Simultaneously, schema knowledge itself, measured with schema generalization on novel similar items, became less precise with time. The authors interpret their results in that schema expression (the influence on episodic retrieval) increases over time as episodic precision decays. Thus, schematic information can be rapidly accessible although it is not necessarily expressed, in line with neurobiological data (Tse at el., 2011).

Based on the presented neurobiological framework, schema-based learning of associations could be subject to different processes than traditional episodic association learning. In the current work we seek to shed light on the functional characteristics of schema-based learning based on this view. In their recent paper, Gilboa & Moscovitch (2021) propose that multiple psychological and neurobiological memory representations are formed, which are at least but not restricted to detail-rich episodic representations, event-specific but detail-poor generic gist representations, integrated schematic representations containing commonalities extracted across episodes, abstracted semantic representations and other representations that are not accessible to consciousness. With time and experience, those representations undergo dynamic changes in strength, stability, and composition. Thereby, the likelihood of their expression (i.e., the influence on episodic memories) changes. Interestingly, the likelihood of expression might be modulated by demands of the retrieval situation and respective retrieval goals (Gilboa & Moscovitch, 2021; Josselyn & Tonegawa, 2020; Tompary et al., 2020). Based on this model and in line with the neurobiological (Tse et al., 2007; 2011) and behavioral evidence (Tompary et al., 2020), it is conceivable that under the influence of a schema, less-specific representations are formed in parallel to episodic representations but their expression (i.e., if they are recruited during retrieval) depends on the retrieval demands and goals (see also Tompary et al., 2020 for a similar argument; see McKenzie et al., 2013; Preston & Eichenbaum, 2013 and van Kesteren et al., 2012 for different views).

In the current work, we assumed that the likelihood of expression of schema representations could be increased when memory is tested with tasks in which episodic information is not necessary for performance, i.e., recognition and implicit memory tasks (for more information, see chapter 2). We seek to shed light on those processes operating on knowledge structures in the learning of novel associations.

2 A functional perspective on schema-based learning in recognition memory

The neurocognitive mechanisms during memory retrieval depend upon the type of memory, which is probed. Over the past century, several such test formats have been used. You could now for example be asked to reproduce all novel compound words you read in the introduction (free recall), or you could be asked to complete *Star* ____ with the word from the introduction (cued recall). Another type of test is recognition memory, i.e., memory for the previous occurrence of an event (Mandler, 1980) in which you are shown the compound word and asked if you recognize it (were *Nannyjournalism* or *Starchair* part of the introduction?[2]). Recognition memory should thus not be seen as a separate memory system but rather as a dimension of the test situation (see Bader, 2014, for a similar argument). Another dimension of the test situation is whether item or associative information is to be remembered.

A typical item memory task requires participants to distinguish old from new, i.e., unpresented, single items, whereby an associative memory task requires judgements about the co-occurrence of items, contexts or features (Yonelinas, 2002) which can be operationalized as distinguishing intact (e.g., *Starchair*) from rearranged pairs like *Nannychair* (e.g., Naveh-Benjamin 2000; Kamp et al., 2017) whereby in a real experiment, not both versions (intact and recombined) of an originally learned compound word would be presented. In the current project, we are particularly interested in associative recognition memory. But how can people "come to make judgements that an item or event has been previously encountered" (Mandler, 1980, p. 252). Regarding this question, two types of models have been proposed. One group of models suggested a single memory process (e.g., Ratcliff et al., 1992) whereas dual-process models (e.g., Mandler, 1980; Yonelinas, 1994) have been argued to better align with empirical data (see Yonelinas & Parks, 2007, for a review).

[2] Whilst *Starchair* was part of the introduction, *Nannyjournalism* is a not yet presented, new compound word.

2.1 Dual-process models of recognition memory

Dual-process models of recognition memory (Mandler, 1980; Yonelinas, 1994; see Yonelinas, 2002; Yonelinas et al., 2010 for reviews) are bestowed upon the assumption that familiarity and recollection make independent contributions to recognition memory and can be separated on functional and neural levels. Mandler (1980) argues that in an extreme case, a recognition process can output a context-free judgement, based on familiarity. However, in the normal course of events, familiarity, which he describes as intra-item integration of perceptual information, and a memory search process occur conjointly. Familiarity is defined as a fast-acting strength process, subjectively perceived as a feeling of knowing (e.g., “I know the word Starchair”; Mandler, 1980; Yonelinas, 2002). Recollection (bearing similarities to Mandler’s search process) denotes the retrieval of specific details of a study event (e.g., “My colleague showed me this interesting article about nannyjournalism”; Yonelinas, 2002; 2010).

Yonelinas (2002) highlights that the episodic-semantic distinction introduced by Tulving (1985) bears similarity to the dual-process model of recognition in that episodic memory leads to conscious remembering what can be assigned to recollection and semantic memory leads to the conscious impression of knowing, which can be seen as the feeling of familiarity. This mapping of processes to systems has also been proposed in a more recent framework by Bastin et al (2019). In addition, memory performance in item recognition tests is thought to rely on familiarity and recollection, whereas associative recognition relying on forming and retrieving arbitrary associations has been ascribed to recollection (Norman & O’Reilly, 2003; Yonelinas, 2002; Yonelinas et al., 2001; 2010).

Next to behavioral evidence on processing speed as well as receiver operating characteristics (ROCs; see Yonelinas, 2002 for a review), and neuropsychological evidence from patients with brain lesions (Düzel et al., 2001; Yonelinas et al., 1997), a dissociation of familiarity and recollection can be shown by means of ERP studies of recognition memory. Those have provided evidence in that ERP components with specific temporal and topographic characteristics are closely linked to familiarity and recollection.

Recollection has a well-established putative ERP correlate, the parietal old/new effect (see Friedman & Johnson, 2000; Rugg & Curran, 2007 for reviews). This ERP effect constitutes more positive waveforms for detected (hits) as compared to correctly rejected items from 500-800 ms after stimulus onset. For language materials, the spatial scalp distribution is pronounced at left-parietal sides (Rugg & Curran, 2007). The validity of the parietal old/new effect has been shown by means of external validation with the remember-know procedure (Gardiner, 1988; Tulving, 1985) in which participants have been instructed to respond *remember* when they can retrieve specific details of the study episode and *know* when their judgement is based on a feeling of familiarity. In combining the remember-know procedure with an ERP approach, Leynes and Bink (2002) found a larger parietal old/new effect for remembered than for known items. The parietal old/new effect is also correlated with the amount of study details retrieved (Vilberg et al., 2006) and absent in patients who cannot recognize by means of recollection (Mecklinger et al., 1998).

As correlate for familiarity-based recognition, an earlier effect preceding the parietal old/new effect has been identified. This so-called FN400 effect has typically a mid-frontal distribution and occurs between 300 and 500ms after stimulus onset. It also takes the form of more positive waveforms for hits than correct rejections. Studies have validated ERP measures of familiarity by showing that experimental manipulations of familiarity modulate the FN400 (Bruett & Leynes, 2015; Bader et al., 2020; see Curran & Rugg, 2007 for a review).

An intriguing question is how the dual-process model of recognition memory can be integrated with the traditional memory systems approaches (e.g., Squire & Zola, 1996). This has been addressed by Henke (2010) in a processing mode model. Henke argues that instead of differing in terms of consciousness, memory systems rather differ in processing operations involved, relying on three variables: If encoding is rapid or slow, association-like or item-like (by means of unitization; *see below*) and if the resulting memory representation is flexible or rigid. Similar to ‘levels-of-processing’ (Craik & Lockhart, 1972) and ‘transfer-appropriate-processing’ (Morris, et al. 1977), the processing mode model highlights that the requirements of the learning situation determine the type of retrieval operation involved and thus, which processing mode is applied (see Henke, 2010). She considers two processing modes by which rapid learning of associations is possible. The first processing mode relies on a learning system located in hippocampus and neocortex. It is capable of rapid encoding of flexible associations. A second processing mode for rapid learning does not rely on the hippocampus but rather on the neocortex and parahippocampal gyrus and includes rapid encoding of unitized associations. Those representations can be triggered when an identical or highly similar cue is presented. This second processing mode can lead to priming (Schacter et al., 2004) and familiarity-based remembering.

2.2 Unitization and unitized representations

Unitization can be understood as a process that results in the formation of a holistic representation out of unrelated stimuli (see Graf & Schachter, 1989) whereby those unrelated stimuli would usually be stored as a flexible association between two or more items (e.g., Henke, 2010). The probability with which an experimental condition promotes unitization is thought of as being continuous. Therefore, unitization as an experimental variable is considered being continuous rather than binary (Parks & Yonelinas, 2015; Yonelinas et al., 2010). Crucially, this continuity assumption does not necessarily pertain the degree towards a single representation is unitized. Rather, this 'levels-of-unitization' (LOU) account takes into consideration methodological boundary conditions. Those boundary conditions allow for inferences about unitization in an experimental condition but not for testing unitization of a single stimulus (Yonelinas et al., 2010).

Unitization has been argued to come with costs (Mayes et al., 2004; 2007; Pilgrim et al., 2012). Here, the integration of two entities to a single one has been shown to hinder access (here familiarity-based retrieval) of the original constituents (Pilgrim et al., 2012) and also results in a rigid memory representation which is not suitable to support recognition of a reversed pair (e.g., Haskins et al., 2008; Wiegand et al., 2010). In addition, unitization has been shown to be more beneficial for associative memory than for item memory of the underling constituents (Parks & Yonelinas, 2015).

It is believed that unitized associations can be feed into processes that are usually restricted to item memory, i.e., they can be retrieved by means of a familiarity process (Jäger et al., 2006; Rhodes & Donaldson, 2007; 2008). Hereby, studies concerned with the neurocognitive mechanisms underlying the retrieval of pre-existing unitized representations, showed that familiarity can contribute to associative recognition for a unitized representation as compared to a condition in which items are relationally encoded. Rhodes and Donaldson (2007) tested associative recognition memory for pre-experimentally associated word pairs (e.g., "traffic-jam"), associated word pairs that are also categorically or functionally related (e.g., "lemon-orange") or word pairs that are only categorically or functionally related (e.g., "cereal-bread"). ERPs recorded during the test phase revealed that only recognition of original associated word pairs ("traffic jam") was accompanied by an early mid-frontal old/new effect, the putative correlate of familiarity. This is in line with the results of a rating study in which participants judged the associated word pairs to be the most unitized ones. The authors interpret the pattern of results in that associated word pairs are perceived as a new entity (i.e., as a compound word, an argument brought forward by Ahmad & Hockley, 2014) which enables their familiarity-based remembering.

An interesting question regards the conditions under which unitization is initiated. Empirical evidence supports the idea that unitized representations can be formed already after one learning instance in an experiment, resulting in costs for recognizing reversed pairs (e.g., Haskins et al., 2008; Wiegand et al., 2010). Encoding conditions initiating a unitization process can be either bottom-up, whereby to-be-learned materials themselves are manipulated to support unitization (e.g., presenting the pair star-chair as either *Star chair* or *Starchair*), or top-down, where unitization is initiated by instructions (Tibon et al., 2014).

In the current work, we focused on top-down unitization. Top-down unitization can be triggered by, for example, interactive imagery (Diana et al., 2011; Rhodes & Donaldson, 2008), imaging an object in the color of a background (Diana et al., 2011), or encoding as a compound word, supported by a congruent definition (Bader et al., 2010; 2014; Haskins et al., 2008; Quamme et al., 2007). We referred to the latter as the definition-sentence paradigm

(Meßmer et al., 2021, chapter 4). As it highly resembles the schema congruency manipulation applied in the current work, research regarding the definition-sentence paradigm will be considered in the following.

In the definition-sentence paradigm (Bader et al., 2010; 2014; Haskins et al., 2008; Quamme et al., 2007), novel word pairs, e.g., SMOKE APPLE are learned either together with an explaining definition or a gap sentence. The definition enables processing the word pair as one entity, i.e., as a compound word (unitization condition). In contrast, the gap sentence, into which the words should be mentally filled-in, enables their processing as a word pair (non-unitization condition).

Example: SMOKE APPLE

A fruit maturing above flames (unitization condition)

The ______ damaged the _______ (non-unitization condition)

During the learning phase, participants are usually instructed to judge how well the definition combines the meaning of the constituents to a compound word (unitization condition) or how well each word can be filled in a sentence gap (non-unitization condition).

On the neuroimaging level, non-unitized arbitrary associations have been found to involve hippocampal processing during retrieval (e.g., Hannula & Ranganath, 2008). However, the formation of unitized representations in the definition-sentence paradigm shows a more pronounced involvement of perirhinal cortex structures and perirhinal activity during encoding was correlated with familiarity during later retrieval (Haskins et al., 2008; but see Bader et al., 2014). As the perirhinal cortex has originally been discussed as neural correlate of item memory (Davachi, 2006), this is in line with the idea that unitized representations are treated like a single item.

Quamme et al., (2007) tested two groups of amnestic patients as well as two healthy control groups in the definition-sentence paradigm. Due to hypoxia, one group of three patients had a specific deficit in recollection-based remembering whilst familiarity-based processing was close to intact. Another group of two patients with lesions in the left medial temporal lobe had equally reduced recollection- and familiarity-based recognition. One control group consisted of seven age- and education-matched participants whilst the other constituted a larger sample of undergraduate students. The definition-sentence-paradigm (see above) was used and after a retention interval of approximately 30 minutes, an associative recognition memory test followed. In this test, participants were shown original (intact) word pairs, rearranged (recombined) word pairs and yet unpresented (new) word pairs and their task was to judge them as old or new, combined with a confidence scale from 1 (confident new or recombined) to 6 (confident old intact). Memory performance did not differ across the unitization and the non-unitization condition for both healthy control groups and the lesion patients with equally reduced familiarity- and recollection-based processing. However, there was a memory advantage for the unitization as compared to the non-unitization condition for the hypoxic patients who can rather rely on familiarity than recollection. This pattern of results indicates that the unitization condition recruits familiarity, even when the unitized representations are formed during the experiment.

In studies on healthy human participants using the definition-sentence paradigm, it has also been shown that unitized representations can be retrieved via familiarity, as indicated by an early old/new effect from 350 to 500 ms after stimulus onset (Bader et al., 2010; Wiegand et al., 2010). In their study, Bader et al. (2010) applied the definition-sentence paradigm in manipulating the conditions (definition/sentence) in a between-subjects manner. In the subsequent test phase, participants were presented with intact, recombined and new (yet unpresented) word pairs and should classify them as intact, recombined or new. Whilst there were no across-condition differences in recognition accuracy, ERPs revealed an early old/new effect in the definition (i.e., unitization) condition but a late parietal old/new effect, indicating recollection, only for non-unitized word pairs in the sentence condition. However, the spatial distribution of the early old/new effect is different from the traditional mid-frontal old/new effect (FN400) which led to the interpretation that newly formed units are recognized by using a different familiarity signal than pre-existing units, indicated by an N400 attenuation effect instead of an FN400 effect (Bader et al., 2010; Mecklinger & Bader, 2020).

In referring to those different familiarity signals, Mecklinger and Bader (2020) propose a neurocognitive account of familiarity in recognition memory, building on the idea that familiarity is not a single construct but rather multiply determined. Hereby, *absolute (baseline) familiarity* can be distinguished from *relative familiarity* (Mandler, 1980). Absolute familiarity can be seen as the strength of the familiarity signal to a fixed point in time which can correspond to word frequency. Relative familiarity describes the increment in familiarity gained during an episode (e.g., an experimental session).

In an illustrative study on single word learning, Bridger and colleagues (2014) electrophysiologically disentangle absolute and relative familiarity. In their experiment, participants were shown high-frequency and low-frequency words they had to read and say aloud. Participants knew their memory would be tested later. During the recognition memory test, they were presented with words shown during learning and new words. They had to classify each word as either old (presented) or new. The experiment included a speeded recognition memory test (with a response deadline) and a non-speeded test. However, the clearest picture on absolute and relative familiarity effects was obtained from the non-speeded test. The results of the non-speeded test show that absolute familiarity, realized by contrasting correctly rejected high-frequency and low-frequency new words, is indicated by an N400 attenuation effect for high-frequency words. This effect is topographically different from a relative familiarity ERP effect (hits minus correct rejections), which is larger for low-frequency than high-frequency words and maps on the FN400 familiarity effect. How these multiple familiarity signals are used in recognition memory has been refined in the neurocognitive framework by Mecklinger & Bader (2020) and is outlined in the following.

2.3 A fluency-attribution account of familiarity in recognition memory

“we argue that “pastness” cannot be found in a memory trace but, rather, reflects an attribution of transfer in performance” – Jacoby et al., 1989, p. 400.

The neurocognitive framework by Mecklinger and Bader (2020) relies on a slightly different approach to familiarity as the one introduced at the beginning of the chapter: Whereby Yonelinas (1994) describes familiarity as a direct consequence of the strength of a memory representation, Jacoby & Dallas (1981) and Jacoby et al. (1989) argue that familiarity is not an inherent item or event characteristic, but arises when relatively fluent processing of an item is attributed to its past experience (see also Bastin et al., 2019; Mecklinger & Bader, 2020). Fluency here refers to "the speed and ease with which a stimulus is processed" (Bastin et al., 2019, p. 2). Fluency can be conceptual (Bader & Mecklinger, 2017; Mecklinger & Bader, 2020) or perceptual in nature and can also affect performance in an implicit memory test (Bastin et al., 2019; Wang, 2019). Whittlesea & Williams (1998; 2000) add another aspect to this view in highlighting that it is not relative fluency per se that matters in recognition memory judgements, but rather the discrepancy between expected and experienced fluency. This add-on could explain the phenomenon why faces of friends do not feel familiar in contrast to faces of strangers (Whittlesea & Williams, 1998): Fluency for faces of friends is expected, not leading to a discrepancy between expected and perceived fluency. If processing a face of a stranger is fluent, however, this is unexpected, leading to a discrepancy that can be attributed to oldness and a feeling of familiarity.

Bastin and colleagues (2019) highlight that the use of fluency signals in recognition is determined by their relevance, as determined by characteristics of the task. Hereby, two factors determine whether fluency cues are used: First, they must be considered to be diagnostic cues for recognition decisions (Westerman et al., 2002) and second, the experienced fluency should be salient in a given context (Jacoby & Dallas, 1981; Westerman, 2008; Whittlesea & Williams, 2001). "People set an internal criterion along the varying dimension of memory strength depending on the task specificities" (Bastin et al., 2019, p. 8). In line with Whittlesea & Williams (1998; 2000), fluency signal strength exceeding the criterion is perceived as surprising (Yonelinas et al., 2010). Of note, in their model, attribution is not restricted to familiarity (Bastin et al., 2019; but see Ionita et al., 2019).

In their neurocognitive account of familiarity, Mecklinger and Bader (2020) add to the fluency-attribution account in arguing that fluency attribution underlies relative familiarity, indicated by the FN400, whilst attribution processes are not assumed to underly absolute familiarity, the latter being indicated by an N400 attenuation effect. They propose that relative familiarity results from the difference between perceived and expected processing fluency, which is perceived as being surprising, and a mnemonic attribution process by means of which this surprising discrepancy in fluency is ascribed to the learning episode. Absolute familiarity can be used in situations in which all stimuli are novel, which makes changes in absolute familiarity diagnostic for prior occurrence. In a recognition memory experiment in which novel compound words are learned with a congruent definition (as e.g., in Bader et al., 2010; 2014; Haskins et al., 2008; Wiegand et al., 2010), the novel word pair is assigned a meaning. This meaning assignment could enable conceptual fluency which – in a situation where all stimuli are novel and should not have a meaning – is diagnostic for recognition. This meaning assignment results in semantic integration of the compound word. This is indicated by an N400 attenuation effect (e.g., Bader et al., 2010; Wiegand et al., 2010).

2.4 Are unitized representations the key to rapid schema-based learning of associations?

In the first chapter, we outlined that schema-based learning of novel associations might constitute a specific learning mechanism, probably relying on partially different neurocognitive processes than traditional associative learning. Further, the mnemonic consequences of schema-based learning might often be overshadowed by traditional episodic memory processes dominating memory performance (Tompary et al., 2020). From this, it follows that recognition memory, relying less on episodic memory processes than recall (Yonelinas, 2002) might be a memory test format that is especially suited to investigate the impact of schema congruency on the learning of novel associations. Here, the consequences of a memory representation formed under the influence of a schema should be observable. In recognition memory, performance can be driven by both, episodic recollection and familiarity.

Referring to the processing-mode model on associative recognition memory (Henke, 2010), rapid learning of novel associations can be achieved in two ways: Either the associations can be fed into a hippocampal process, resulting in flexible memory representations, or in another non-hippocampal process, resulting in rigid representations that can be recruited by means of a familiarity process when the underlying representations are unitized. In combining the idea that schema-congruency initiates a specific learning mechanism relying less on hippocampal contribution and that associations can be learned with less hippocampal contribution when unitized representations are formed, we hypothesized that schema-congruency might support the formation of unitized representations. Hereby, the existence of a unitized representation resulting from semantic integration of the compound word might result in conceptual fluency, which is diagnostic in a recognition memory test.

Interestingly, it has been assumed that unitization requires an entity-defining framework (Bader et al., 2010; Mecklinger & Jäger, 2009) which we argue could be provided by schema-congruency, which is established via the congruent definition. First evidence for a role of schemata in the definition condition comes from the study by Bader and colleagues (2014) who found recognition of novel units in the definition, but not in the sentence condition being accompanied by activity in the mPFC, which is strongly associated with schema-based memory (see van Kesteren et al., 2012).

In the current work, we sought to shed light on the neurocognitive processes underlying schema-based learning and recognition of novel word associations, i.e., novel compound words.

In the current series of experiments, schema-congruency was operationalized as how well a congruent definition explains the contribution of both compound words' constituents to a novel concept. The congruent condition resembled the definition condition in the definition-sentence-paradigm (Bader et al., 2010; 2014; Haskins et al., 2008; Quamme et al., 2007; Wiegand et al., 2010). Here, the contribution of both constituents to the whole-word meaning of the novel compound word was explained. In the schema-neutral control condition, only the contribution of the second constituent was explained, whereby no conceptual support was provided for the first constituent of the compound word. Thus, no congruent schema could be used to process the novel compound word as a whole new concept.

Example:

Ein Lexikon, das Gärtner$_{\text{Congruent}}$/ Lehrer$_{\text{Neutral}}$ benutzen, heißt... Gemüsebibel

A dictionary used by gardeners$_{Congruent}$/ *teachers*$_{Neutral}$ *is called... vegetable-bible*

The schema-congruency manipulation was applied in all three experiments. In an incidental learning phase, compound words were presented, preceded by either a congruent or a neutral definition. Participants should rate how well the novel compound word is explained by the preceding definition. After a short retention interval of several minutes, memory was tested in a recognition test (experiments 1 and 3) or in an implicit lexical decision task (experiment 2).
The following three research questions (RQs) were addressed:

2.4.1 RQ 1: What are the neurocognitive mechanisms driving schema-based learning and recognition of novel compound words?

In a first step, we were interested in the neurocognitive processes underlying the encoding of novel associations under the influence of a congruent schema. This question was addressed in experiment 1. Hereby, novel compound words were incidentally learned with either a congruent or a neutral definition. In a later recognition memory test, participants were shown original (intact) compound words, recombined compound words and new (yet unpresented compound words) and should decide for each compound words if it was intact, recombined or new. To shed light on the neurocognitive processes involved in schema-based encoding of novel associations, ERPs were recorded during learning and analyzed with the subsequent memory approach (Paller et al., 1987; Paller & Wagner, 2002). Hereby, we were interested in the cognitive encoding processes underlying successful memory performance in a subsequent recognition memory test.

2.4.2 RQ 2: What type of representation is formed and what are the neurocognitive processes involved during their recognition?

We hypothesized that schema-congruency might enable the formation of unitized representations of novel compound words. We tested this hypothesis in experiment 1. Hereby, we analyzed ERPs recorded during the associative recognition memory test. We expected a larger N400 attenuation effect for compound words initially encoded in the schema congruent as compared to the neutral condition, indicating a larger contribution of absolute familiarity to recognition memory. Whilst this pattern of ERP results would indicate more unitization in the congruent, as compared to the neutral condition, we were additionally interested in the behavioral data pattern that associative memory performance should benefit more from unitization than item memory performance (Parks & Yonelinas, 2015).

Next to the question whether schema-based learning enables the formation of unitized representations, we also sought to investigate whether those representations are less detailed, resulting in more false alarms to semantically similar compound words as e.g., Gemüsebibel (Vegetablebible) when the corresponding original compound word is Gemüseglossar (Vegetableglossary). This question was addressed in experiment 3. Therefore, the recognition memory test from experiment 1 was modified and instead of new compound words, semantically related lure compound words were presented.

2.4.3 RQ 3: How can these representations be retrieved?

Henke (2010) argued that consciousness does not play a role for how memory representations are used (see also Reder et al., 2009; Züst et al., 2019). In her model, unitized representations can be used within a familiarity process but should also result to priming effects. Thus, we investigated whether schema-congruency-driven unitization affects performance in a lexical decision task (experiment 2). In this task, intact, recombined and new (not yet presented) novel compound words were presented together with real compound words and participants should classify each word as real or not real. Here, we assumed that semantic integration of novel compound words, supported by schema congruency and resulting in a unitized representation, should result in conceptual fluency, which is misused in lexical decisions, leading to more errors on novel compound words learned under the influence of a congruent schema than on compound words from the schema-neutral condition. Further, if this phenomenon is associative in nature, there should be a larger associative priming effect in the schema-congruent than in the schema-neutral condition.

3 Stimulus materials

To address the research questions presented in chapter 2, we created a large pool of stimulus materials that are used in all three experiments presented in this thesis. Therefore, we set several requirements the stimulus materials must fulfil to be appropriate for the current series of experiments. The stimulus materials are available from https://osf.io/d7w5a/

Firstly, we want to investigate schema congruency in a highly ecologically valid language setting. As argued by Maguire (2022) and Ranganath (2022), memory research should adopt a focus on more natural processing situations. Although our stimulus materials still are highly controlled, they can be seen as a step towards more rich and thereby natural stimuli. Thus, the definitions that should activate schema knowledge were formulated as complete sentences and carefully screened from several researchers and student research assistants from the department to be well-formed German sentences. In addition, the compound words were created with the idea that the concepts represented by the compound word could be a yet unknown concept. The constituents of the compound words must be pre-experimentally unrelated to avoid interference with existing semantic relationships. Secondly, the schema congruency manipulation should be tightly controlled, especially in terms of language properties. Thus, the schema congruent and the neutral definition are only allowed to differ in one single noun. Schema congruency was induced by a specific pattern of semantic relationships (see 3.1.).

Our initial goal was to obtain 240 compound words, each with both a congruent and a neutral definition, together with associated recombined (experiment 1, 2 and 3) and lure compound words (experiment 3) that were required for the different types of memory tests. In addition, 80 new compound words were required for the recognition memory test in experiment 1. For experiment 2, 140 stimuli were selected from the 240 stimuli used in experiment 1 and 3 (see 3.3. for a precise description).

3.1 Logic of stimulus creation

In a first step, we created 300 novel noun-noun compound words, adapted from Bader et al. (2014). Each compound word consists of two pre-experimentally unrelated nouns. Whenever necessary, the nouns were modified to create a grammatically legal and content-wise plausible compound word. Therefore, interfixes (-s, -n, -en) were included and some nouns appear in plural. The plural form is made plausible by the wording of the definition. The compound words were endocentric and subordinate in that the first word was always the modifier, and the second word was always the head.

For each novel compound word, both a congruent and a neutral definition were created. A definition was stated congruent when it reasonably explains the novel combination of the two nouns to a novel compound word, including a new concept.

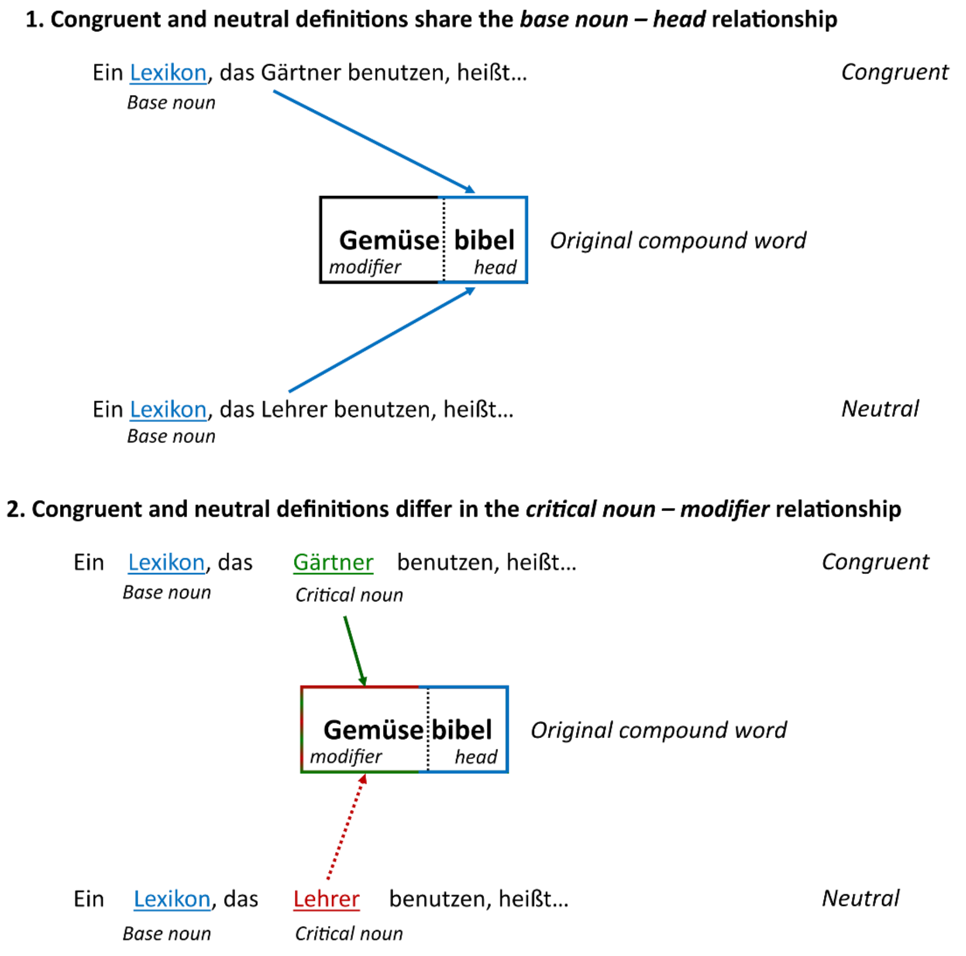
Example:
Ein Lexikon, das Gärtner$_{\text{Congruent}}$/ Lehrer$_{\text{Neutral}}$ benutzen, heißt... Gemüsebibel
A dictionary used by gardeners$_{\textit{Congruent}}$*/ teachers*$_{\textit{Neutral}}$ *is called... Vegetablebible*

This was achieved by using a systematic pattern of relationships between compound head, compound modifier and particular words of the definition sentence as construction principles.

Firstly, the main component of the definitions was a noun phrase, including a single noun (the base noun; “dictionary”) and a relative clause (“used by gardeners”). The base noun bore a semantic relationship to the head of the novel compound word (“bible”) and thus established a link to the core concept. Secondly, the relative clause further specified the established concept. Congruent and neutral definitions only differed in one single noun in the relative clause, i.e., the critical noun (“gardeners”) in the congruent definition, and “teachers” in the neutral definition). In the congruent definition, the critical noun (“gardeners”) bore a semantic relationship to the modifier of the novel compound word (“vegetable”). Thus, the congruent definition combined the underlying concepts of modifier and head to a coherent concept. In the case of the neutral definition, the critical noun (“teachers”) was semantically unrelated to the modifier of the novel compound word (“vegetable”). Consequently, in the neutral condition, the underlying concepts of modifier and head were not combined to a coherent concept (See Figure 1 for an overview of the semantic relationships).

Figure 1

Overview over the semantic relationships between the definitions and the respective original compound word.



Note. The first panel shows the shared semantic relationship between the base noun and the head of the compound word. The second panel shows the semantic relationship between the critical noun and the modifier of the compound word, which is present in the congruent condition and absent in the neutral condition.

Definitions always contained 5–12 words and were completed by two German versions of the formulation "is called" (*heißt* and *nennt man*), to establish a more natural processing situation. Those were used in a way that did not require grammatical alternations of the sentences. For both sets used in experiments 1 and 3 as well as for the seven learn lists in experiment 2, each formulation occurred in approximately half of the trials. The three dots following the definition in the example are for illustrative purposes, only, and were not shown in the experiments.

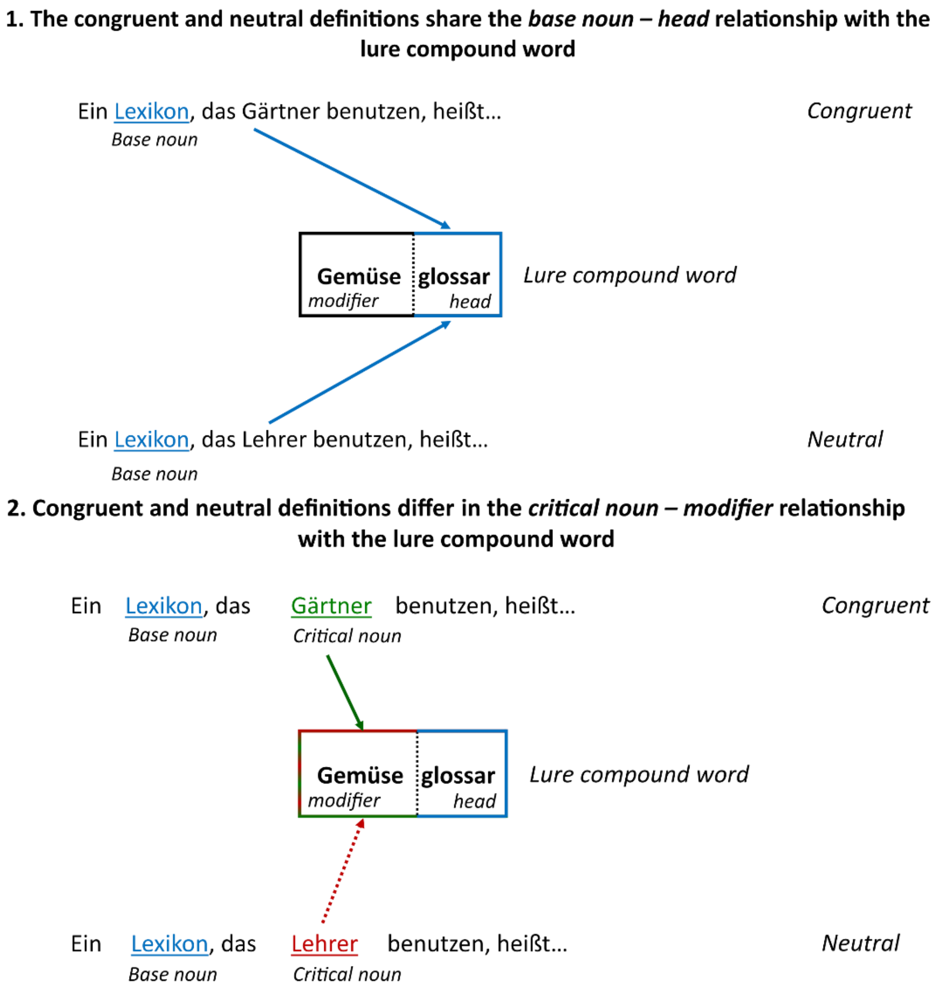
For the memory tests, we additionally created 150 recombined compound words, 300 lure compound words, as well as 118 new compound words. New compound words consisted of two unrelated nouns, which were not used elsewhere in the material (e.g., *Ankermönch* anchor-monk, *Damenraster* lady-grid).

Recombined compound words were included to assure that participants would not be able to solve the task by using item recognition alone and constructed by newly combining the modifier and the head of two different compound words. It was assured that the nouns still were semantically unrelated and for each such pair of compound words, only one of two possible recombined compound words was used, to avoid the repetition of the constituents in different recombined word pairs. For example, we recombined the word *Sternensessel* (star-chair) with the word *Magnetenozean* (magnet-ocean) to *Magnetensessel* (magnet-chair) and omitted *Sternenozean* (star-ocean) from the test list. The two compounds used for recombination were always of the same grammatical gender and contained the same type of interfix, if any.

For the third experiment presented in chapter 7, we additionally created semantical lure compound words. Therefore, the original compound word was modified by exchanging the head with a semantically related word, e.g., *Gemüsebibel* (Vegetablebible) was modified to *Gemüseglossar* (Vegetableglossary; see Figure 2 for an overview of the semantic relationships between definitions and the lure compound word).

Figure 2

Overview over the semantic relationships between the definitions and the respective lure compound word.



Note. The first panel shows the shared semantic relationship between the base noun and the head of the lure compound word. The second panel shows the semantic relationship between the critical noun and the modifier of the compound word, which is present in the congruent condition and absent in the neutral condition.

3.2 Stimulus material selection for experiments 1 and 3

To evaluate the appropriateness of the stimulus materials, three rating studies were conducted. In the first rating study, we verified that the constituents comprising the compound word, the recombined, the new and the lure compound word were pre-experimentally semantically unrelated. In addition, we evaluated the pattern of semantic relationships between the compound word and the definition described above.

In the second rating study, we validated the critical schema congruency manipulation. The final stimulus materials used in all three experiments were selected based on the results of these two rating studies. In the third rating study, we ruled out an alternative explanation for schema congruency effects, which is that possible experimental effects result from differences in imageability of the congruent and the neutral definition.

3.2.1 Methods rating study 1: Semantic relationships.

The described pattern of semantic relatedness and unrelatedness was evaluated in a first rating study (see Figures 1 and 2). An important requirement for the stimulus materials was that the nouns forming novel compound words, lure compound words, recombined compound words and new compound words should be semantically unrelated. The second assumption was that the head of the novel compound word (Bibel, *bible*) and the head of the lure compound word (Glossar, *glossary*) are semantically related. Lastly, the described pattern of semantic relationships between the novel and lure compound word and the congruent and neutral context was to be evaluated.

Regarding the systematic pattern of semantic relationships between the definition and the compound word, we assumed that the *basis noun* (Lexikon, *dictionary*) and the head of the compound word (Bibel, *bible*), as well as the *basis noun* and the head of the lure compound word (Glossar, *glossary*) are semantically related. Lastly, the modifier of the compound word (and the lure compound word), i.e., Gemüse (*vegetable*) and the *critical noun* (Gärtner, *gardeners*) in the congruent definition were assumed to be semantically related whereas the modifier and the *critical noun* in the neutral definition (Lehrer, *teachers*) were assumed to be unrelated (see Figure 1 and 2).

Therefore, the stimulus materials were divided into two subsets, each consisting of 150 compounds, together with their corresponding lure compound word, their congruent and neutral definition as well as their recombined compound word. The two compound words belonging to a common recombined compound word were always assigned to a set together. The list of new compound words was also divided into two halves and one half was assigned to one set. For each set, the word pairs forming the to-be-evaluated relationships described above were extracted. Each set thus consisted of 1184 word pairs (600 which should be related and 584 which should be unrelated).

A sample of *N* = 25 participants (one twice) took part in this rating study. One participant was excluded from the analyses due to a crash of the computer program. Another participant took part twice in the study and therefore, the second data set was excluded. This resulted in a final sample of *n* = 24 participants (15 female, 9 male, *Mdn* = 23 years, *SD* = 3.78, age range = 18-32), half of them rating the first stimulus set and half of them rating the second stimulus set.

The rating study was conducted in E Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Participants were instructed to decide how related the presented words are, concerning their meanings. Hereby, relatedness was described rather broadly, including relatedness due to describing the same concept, shared features of the underlying concepts (categorical relationship) or frequent co-occurrence (thematic relationships). They were asked to use the scale 1 (not at all), 2 (rather not), 3 (rather), 4 (absolutely) by using the keys d, f, j and k on the keyboard. Key assignment (ascending or descending scale) and set assignment was counterbalanced across participants. Each trial started with a fixation cross, which was presented for 500 ms, followed by the stimulus display. On the stimulus display, the two nouns were presented with 4 spaces in between. Below the nouns, the question and the scale were presented. The stimulus remained on the screen until participants responded, but for a maximum of 5000ms. After, a short blank of 10ms followed before the next trial started. The order of the nouns within the word pair was always the same as in the final EEG experiment with the noun from the definition always coming first and all types of compound words presented in their natural arrangement.

Stimulus presentation order was pseudorandomized with the restriction of not more than three consecutive trials displaying a related or unrelated word pair, respectively. The semantic relationship within each noun word pair was evaluated by *n* = 12 participants. Due to response deadline, there were missing values in the ratings on semantical relatedness. However, a minimum of 11 ratings were available for each word pair.

3.2.2 Methods rating study 2: Appropriateness of the definition.

In a second rating study, it was to be evaluated how well the combination of the two words in the novel compound word denotes the concept, given by the definition. Therefore, we tested *N* = 32 additional participants (21 female, 11 male, *Mdn = 23* years, *SD* = 3.66, *age range* = 18-31).

The rating study was conducted in E Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Each participant rated all 300 novel compound words, half of them with the congruent definition and the other half with the neutral definition. This assignment was counterbalanced across participants. Participants were asked to indicate how well the combination of the two words denotes the concept, which is given by the definition on a four-point scale (1 = *not at all*, 2 = *rather not*, 3 = *rather*, 4 = *absolutely*) by using the keys d, f, j and k on the keyboard. Key assignment (ascending or descending scale) was counterbalanced across participants. Each trial started with a fixation cross, which was presented for 500ms, followed by the stimulus display. On the stimulus display, the definition was presented with the compound word underneath. Below the compound word, the scale and the question were presented. The stimulus remained on the screen until participants responded, but for a maximum of 10000ms. After, a short blank of 10ms followed before the next trial started.

3.2.3 Results rating studies 1 & 2 and stimulus selection.

The final stimulus materials were chosen based on the ratings in both rating studies, with the rule that novel compound words, which form a recombined compound word together, were included or excluded pairwise. Selection and analyses were done in R (version 3.6.1; R Core Team, 2019) and RStudio (Version 1.2.5001; RStudio Team, 2019) as well as IBM SPSS statistics (version 26).

The following criteria were used: First, due to spelling errors, 16 stimuli were excluded. Second, the novel compound word, the recombined compound word and the lure compound word were required to be unrelated, i.e., have a rating value from smaller or equal to 2 in the first rating study. Therefore, an additional 32 stimuli were excluded. In the next step, the difference in the ratings of the congruent and the neutral context gathered in the second rating study was required to be a positive value. This criterion was set to exclude combinations in which the neutral definition explains the novel compound word better than the congruent definition. Therefore, 6 additional stimuli were excluded. Lastly, the best 240 stimuli were selected from the remaining 254 based on the largest difference in rating of the congruent and the neutral definition. Table 1 shows the rating results for the nouns within all types of compound words. In line with our requirements, all types of compound words were rated low in semantic relatedness, i.e., close to 1. Figures 3 and 4 display the rating results for the semantic relationships between the definitions and the original, as well as the lure compound word for this stimulus selection.

Tabelle 1:

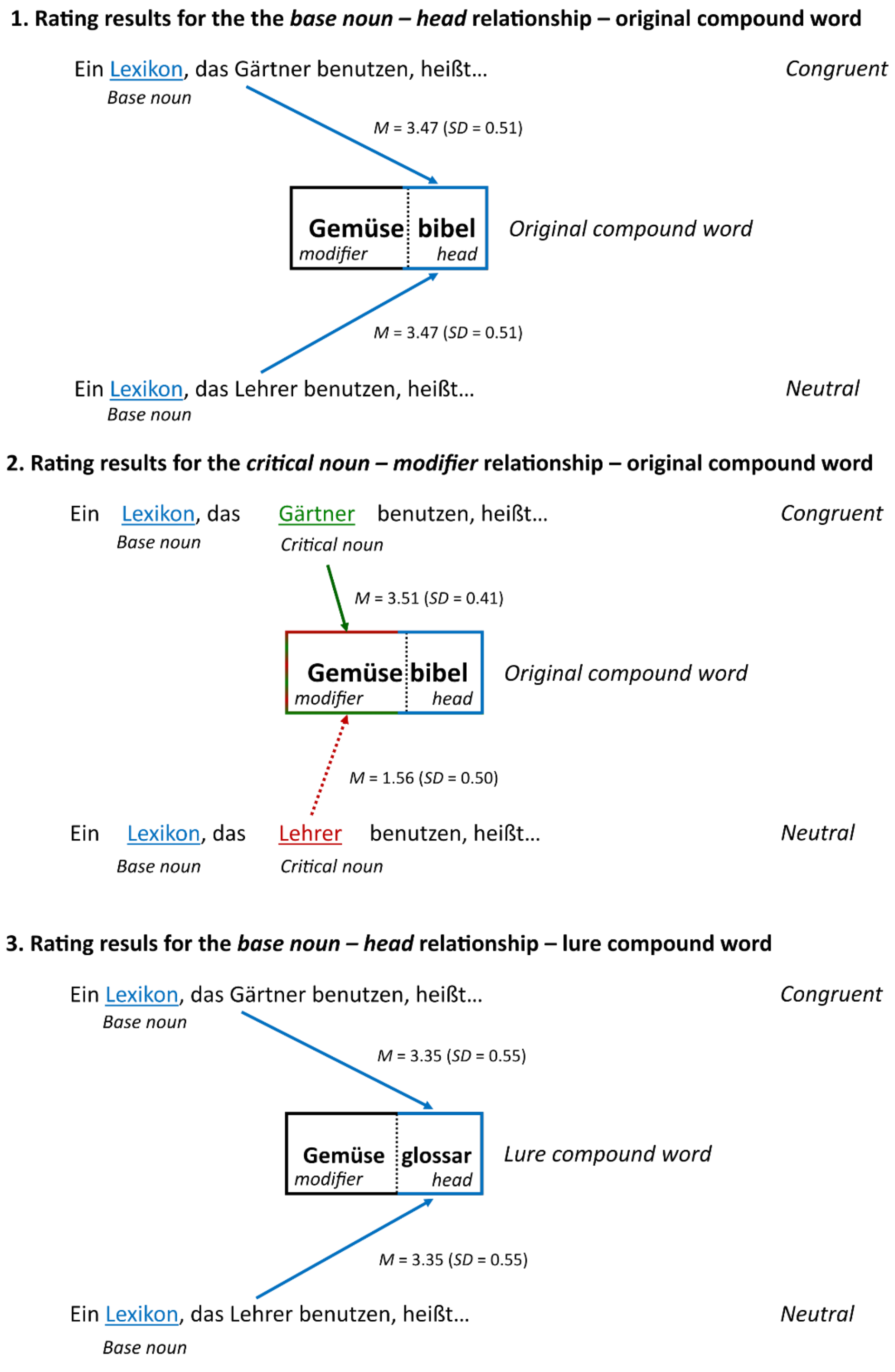
Results of the first rating study on semantic relatedness for the compound word constituents

Compound word type	**Rating value**
Original (e.g., Sternensessel)	M = 1.30, SD = 0.23
Recombined (e.g., Magnetenozean)	M = 1.28, SD = 0.22
New (e.g., Ankermönch)	M = 1.30, SD = 0.14
Lure (e.g., Sternenstuhl)	M = 1.31, SD = 0.23

Note. Rating results are presented for the final stimulus material selection.

Figure 3

Overview over the rating results regarding the semantic relationships.

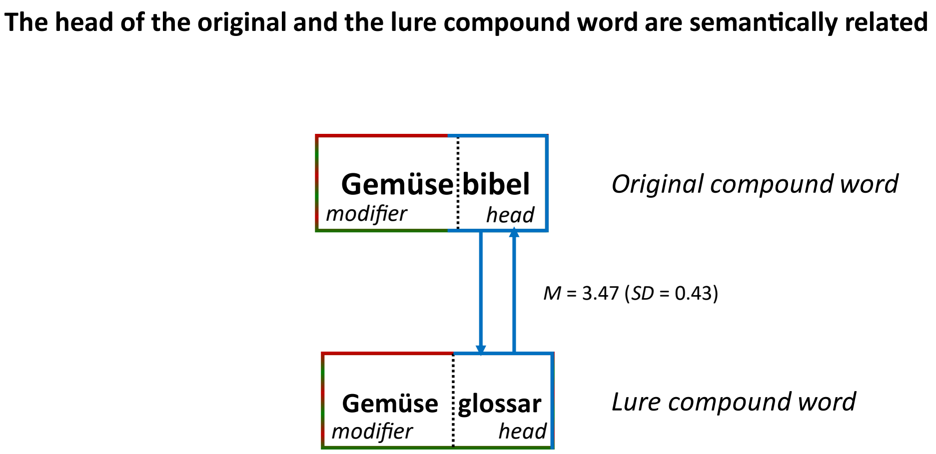


Note. The first panel shows the results for the shared semantic relationship between the base noun and the head of the original compound word. The second panel shows the semantic relationship between the critical noun and the modifier of the original compound word, which is present in the congruent condition and absent in the neutral condition. The third panel shows the results for the shared semantic relationship between the base noun and the head of the lure compound word. The rating values were accessed a scale from 1 (not at all) to 4 (absolutely).

Those confirm the pattern of semantic relationships we used to achieve our critical schema congruency manipulation: Firstly, the *head noun* of both, the original and the lure compound word is semantically related, i.e., rated close to 4, to the *base noun* of both the congruent and the neutral definition. Secondly, the modifier noun of both, the original and the lure compound is semantically related to the critical noun in the congruent definition, but less related to the critical noun in the neutral definition.

Figure 4

Rating results for the semantic relationship between the head of the original and the lure compound word.



Note. The rating values were accessed a scale from 1 (not at all) to 4 (absolutely).

Lastly, the head of the lure compound word and the original compound word were rated as semantically related, as it would be expected for a semantically related lure.

Regarding the appropriateness of the definition, the congruent definition was rated as significantly more explaining the novel compound word than the neutral definition ($M_{congruent}$ = 3.16, *SD* = 0.40, $M_{neutral}$ = 1.69, *SD* = 0.36, $t(239) = 42.99$, $p < .001$, $g_{av} = 3.82$)[3].

[3] Note that due to the time restriction, there were missing values in rating responses. However, a minimum of 11 ratings were available for each version (congruent and neutral) of the definition.

Compound words had a length ranging from 7 to 18 letters (*M* = 11.98, *SD* = 2.09), for to-be-learned compound words, ranging from 7 to 16 letters (*M* = 11.93, *SD* = 2.03) for recombined compound words, ranging from 8 to 20 letters for lure compound words (*M* = 12.35, *SD* = 2.29) and ranging from 7 to 18 letters for new compound words (*M* = 12.37, *SD* = 2.58). Normalized lemma frequency was estimated from dlexDB (Heister et al., 2011). For one noun, there was no respective entry. The normalized lemma frequency for the constituents ranged from 0.02 to 264.38 occurrences per million for compound words and recombined compound words, from 0.01 to 233.63 occurrences per million for the second constituent of the lure compound words and from 0.11 to 394.69 for new compound words.

3.2.4 Rating study 3: Imageability.

It could be argued that the two definition types do not only differ in the semantic congruency between the definition and the compound word, but also in the amount of semantic content of the definitions themselves. Congruent contexts may have been richer in content and thus may allow better prediction of the target words (see Federmeier et al., 2007, as an example). To test whether both context types differ in their predictability, we conducted an additional rating study in which we presented congruent and neutral definitions without the compound word and asked participants to indicate how well they could imagine something from the definition. The rationale behind this approach was to check if the congruent condition induces more constraint and thus predictive potential than the neutral condition, when presented without the compound word. As a cloze study is not suitable to estimate constraint for novel compound words, constraint was operationalized as imageability, i.e., how well someone could imagine something from the definition.

The rating study was conducted online, generated using SoSci Survey (Leiner, 2019) and was made available to users via https://s2survey.net . A sample of *N* = 27 participants took part in the study. Therefore, two lists were created in which 120 fictional definitions were presented in their congruent version and the remaining 120 fictional definitions in their neutral version. The assignment of the congruent or neutral version of the definition was counterbalanced across lists and list assignment was counterbalanced across participants so that each definition in each version was rated by $n = 13$ or $n = 14$ participants, respectively. Participants saw the sentences without “is called”, but instead completed with three dots (e.g., “Ein Lexikon, das Gärtner benutzen...”, *A dictionary used by gardeners ...*) and had to indicate how well they could imagine something from the definition on a four-point scale (1 = *not at all*, 2 = *rather not*, 3 = *rather*, 4 = *absolutely*). No time limit was given, and items could not be skipped. There was no significant difference between conditions for these ratings, $t(239) = 1.90$, $p = .058$, $g_{av} = 0.15$ ($M_{congruent} = 2.97$, $SD = 0.47$, $M_{neutral} = 3.04$, $SD = 0.45$), with an opposite numerical trend speaking against the idea that the contexts differ in imagineability.

3.3 Stimulus material selection for experiment 2

For the adapted lexical decision task (Parks & Yonelinas, 2015) in experiment 2, we required real compound words to-be-distinguished from our novel compound words to make the task plausible. Thus, 100 real compound words were selected from the materials provided by Schulte im Walde & Borgwaldt (2015) and Bader (2014). To ensure that real compound words and novel compound words do not share any constituent nouns and also selecting a similar number of stimuli as Parks & Yonelinas (2015), another stimulus selection was made for the novel compound words. Hereby, a subset of 140 novel compound words was selected out of the 240-compound words used in experiments 1 and 2. For those compound words and their respective definitions, no nouns were shared with the constituents of the real compound words. Novel compound words had a length ranging from 7 to 18 letters for the original compound words ($M = 12.00$, $SD = 1.97$) and ranging from 7 to 16 letters for recombined compound words ($M = 12.01$, $SD = 1.91$). Note that the new compound words presented in this experiment are a selection of the original compound words. The real compound words had a length ranging from 9 to 19 words ($M = 11.73$, $SD = 2.23$).

The normalized lemma frequency for the constituents of the compound words ranged from 0.02 to 229.47 occurrences per million for original compound words and from 0.02 to 224.45 occurrences per million for the recombined compound words. The real compound words had a normalized lemma frequency ranging from 0.01 to 10.96 occurrences per million. For 6 compound words, there was no respective entry.

3.4 Semantic transparency

In a fourth rating study, we addressed the idea that our schema congruency manipulation induced differences in semantic transparency of the compound words which we would expect based on the similarity of the concepts.

3.4.1 Rating study 4: Semantic transparency.

Interestingly, the linguistic concept *semantic transparency* resembles schema congruency as it is operationalized in the current thesis. Semantic transparency can be defined as how transparent the end product of a morphological process is with regard to its meaning (Bell & Schäfer, 2016). In a word like *Tea cup*, the contribution of both constituents to the whole word meaning is clear whereby this is not the case for *elderberry*. Similar to congruency with a schema, semantic transparency is known to play a role in the processing and also storage of compound words (Marslen-Wilson et al., 1994).

In the current project, the congruent definition explains the contribution of each constituent to the meaning of the compound word and thus leads to a fully transparent compound word. In contrast, the neutral definition only explains the contribution of one word to the meaning of the compound word, resulting in a half-transparent compound word. Thus, the first hypothesis was that semantic transparency should be rated higher in the congruent, as compared to the neutral condition. Secondly, given the high similarity of schema congruency and semantic transparency, we expected a high correlation between the definition fit ratings from rating study two and the current rating study.

This hypothesis was evaluated in an additional rating study. This rating study was conducted online, generated using SoSci Survey (Leiner, 2019) and was made available to users via https://s2survey.net .

A sample of N = 25 participants took part in the study. Therefore, two lists were created in which 120 fictional definitions were presented in their congruent version and the remaining 120 fictional definitions in their neutral version. The assignment of the congruent or neutral version of the definition was counterbalanced across lists and list assignment was counterbalanced across participants so that each definition in each version was rated by $n = 12$ or $n = 13$ participants. Participants saw the definitions without the compound words and were asked how well they could imagine something from the definition on a four-point scale (1 = *not at all*, 2 = *rather not*, 3 = *rather*, 4 = *absolutely*). No time limit was given, and items could not be skipped. Due to a list error, data for 5 compound words could not be analyzed. There was a significant difference between conditions for these ratings, $t(234) = 42.29$, $p < .001$, $g_{av} = 3.68$ ($M_{congruent} = 3.16$, $SD = 0.42$, $M_{neutral} = 1.70$, $SD = 0.37$, as it was for the context fit rating for this stimulus subset, $t(234) = 41.53$, $p < .001$, $g_{av} = 3.82$ ($M_{congruent} = 3.16$, $SD = 0.40$, $M_{neutral} = 1.69$, $SD = 0.36$). The transparency rating was highly correlated with the context fit rating for both the congruent condition, $r = .649$, p < .001, and the neutral condition, $r = .723$, $p < .001$, supporting the idea that schema congruency might enable semantically transparent processing of novel compound words.

4 Schema-congruency supports associative encoding of novel compound words (Exp. 1).

Experiment 1 spans over chapter one and two, mapping on two different research questions addressed, which are RQ 1 and RQ 2 (see chapters 2.4.1 and 2.4.2, respectively). In the current chapter, the influence of schema-based learning on neurocognitive processes during encoding is covered[4].

4.1 Introduction

Imagine you are reading a newspaper article. Eventually, you stumble over the word flight shame. Up to now, you do not know what flight shame means, but by reading the article, you learn that it denotes feelings of shame about flying, due to its negative consequences for the environment. This definition provides a plausible explanation for the combination of the words flight and shame to a novel concept. Now imagine a similar scenario, but this time, you come across the word acrophobia. While reading the article, you learn that this phobia is about the fear of heights. This time, it is much harder to track the contribution of each constituent to the novel concept, as you do not know any Greek and thus you cannot make sense of the first constituent. In the first case, you can integrate the novel concept into your prior world knowledge, as you know the underlying concepts. However, this is not possible in the second case, in which you do not understand the contribution of *acro* to the meaning of the word and thus may not integrate this constituent into your prior knowledge structure. An interesting question is how this congruency with prior knowledge influences memory formation of novel compound words, for which a novel concept is created, i.e., the episodic encoding of two previously unrelated items (constituents) into an associative memory representation.

It is well established that events which are congruent with a given schema are better retained than schema-incongruent events (Alba & Hasher, 1983; Schulman, 1974; Pichert & Anderson, 1977; see Greve et al., 2019, for an

[4] This chapter consists of a modified version of a published manuscript (Meßmer et al., 2021) in which paragraphs have been removed or slightly edited to avoid redundancy.

overview), although up to now, only few studies have investigated the influence of a memory schema on the learning of associations.

Staresina et al. (2009), for example, operationalized event congruency as the plausibility judgment given by participants for the semantic match of a word–color combination. They found superior item memory and also better memory for associated source details (i.e., color information) for congruent events. Bein et al. (2014) presented semantically and associatively related and unrelated word pairs at study and found substantially elevated memory scores for item and associative memory for related (schema supported) words. However, an intriguing question is by means of which processes a schema supports memory formation and retrieval in general and for associations in particular.

On the functional level, the memory advantage for schema-congruent events has been ascribed to easier integration of information that matches representation in semantic memory. This leads to richer and more elaborated memory traces, which are more accessible in a subsequent memory test (Craik & Tulving, 1975; see also Bein et al., 2015). However, it remains to be specified how exactly the presence of a schema supports episodic encoding and leads to beneficial effects on subsequent recognition and recall. Thus, a primary goal of the present study was to assess the mechanisms by which prior semantic schema knowledge facilitates episodic encoding of two unrelated items into an associative memory representation.

An ERP measure that can be used as an index of semantic processing is the N400. During natural reading, the N400 is attenuated for words that are semantically congruent with a preceding context (Kutas & Hillyard, 1980). This is referred to as the N400 effect. Based on a large number of studies, the N400 has been linked to the retrieval and integration of semantic information (see Kutas & Federmeier, 2011, for a review). Of particular interest in the present study was whether the facilitated processing of schema-congruent events is reflected in an attenuation of the N400 (semantic priming effects), and whether this N400 attenuation effect is predictive for subsequent memory of these events.

To explore the mnemonic processes involved in schema-based learning, we used an ERP measure that is indicative of successful memory encoding, the subsequent memory effect (SME), or Difference in neural activity due to memory (Dm effect; Paller et al., 1987; Paller & Wagner, 2002). An SME is obtained by comparing ERPs during the encoding of events that are remembered versus forgotten in a subsequent memory test (Sanquist et al., 1980; see Cohen et al., 2015; Paller & Wagner, 2002, for reviews). Thereby, SMEs serve as online measures, reflecting processes that are associated with later successful memory performance.

Packard et al. (2017; Exp. 4) applied the SME logic in a schema-based learning study by using a variation of the Deese-Roediger-McDermott (DRM) task (Roediger & McDermott, 1995). In this task, participants were presented with word pairs of category labels and exemplars, the latter being either congruent or incongruent with the category label. In a subsequent memory test, participants were again presented with these exemplars, together with new words and had to decide if the presented word was old (presented in the study phase) or new. Packard et al. (2017) found SMEs for both, semantically congruent and incongruent words. However, these effects unfolded around 200 ms earlier for congruent than incongruent words, which led the authors to conclude that semantic congruency accelerates episodic memory encoding. This interpretation, however, was recently challenged because in the first experiment of their paper, not only congruent words were more often classified as old, but also semantic lures. Consequently, rather than promoting episodic encoding, there might have been a semantic bias, boosting old responses to all exemplars congruent with a studied category. As no semantic lures were present in the experiment in which the SMEs were reported, it cannot be ruled out that the earlier onsetting SME effect in the congruent condition is an N400 effect, reflecting the facilitated semantic access to items that are semantically related to the target word, rather than the successful episodic encoding of schema-congruent events (Höltje et al., 2019; Tibon et al., 2017).

A recent study from our lab (Höltje et al., 2019) further explored this issue and compared SMEs and memory performance for words which were either congruent ("dog") or incongruent ("sapphire") with a preceding category phrase ("a four-footed animal"). In a surprise subsequent memory test, participants had to discriminate studied (old) words from semantically related lures, i.e., words fitting to a studied category phrase, but not presented during the learning phase ("fox"). Memory was better for congruent words and, in contrast to the results by Packard et al. (2017), Höltje et al. did not find any temporal differences between the SME to congruent and incongruent words. Rather, an SME from 300 to 700 ms with a parietal topographic maximum was obtained for congruent words, only. Notably, this parietal SME can be traced back to successful schema-supported episodic encoding and sheds light on the processes involved in the schema-congruency effect. As similar effects have primarily been found in memory tasks that probe memory for single items (Fabiani et al., 1986), this SME has been linked to the processing of item-specific details, possibly increasing the distinctiveness of an item (Fabiani et al., 1986; Höltje et al., 2019; Karis et al., 1984). In any event, this finding suggests that schemas support memory formation by enhancing the formation of item-specific details, or by integrating new information with pre-existing knowledge.

However, our initial question remains: How does schema knowledge facilitate episodic encoding of two unrelated items into an associative memory representation and the creation of a novel concept? A first hint concerning this question is provided by a recent study by Kamp et al. (2017), in which the learning of associations was investigated in a unitization task. Unitization refers to a condition in which previously separate items are integrated and become represented as a new single unit (Graf & Schacter, 1989). Kamp et al. (2017) presented unrelated word pairs together with a definition, which fosters the processing of those words as compound words with a new joint meaning, i. e., a novel concept (enabling unitization encoding). In a control condition, the two words had to be filled in a sentence, resulting in their processing as separate items. We argue that the definition condition fosters schema-based encoding of the word pair, whereas this form of encoding is largely absent in the sentence condition. Interestingly, Kamp et al. (2017) found a subsequent memory effect with a parietal maximum resembling the SME in the study by Höltje et al. (2019). This effect was present in the definition condition and virtually absent in the control condition. The authors interpreted it as reflecting the encoding of rich item-specific details of the to-be-encoded unit. However, the Kamp et al (2017) study was not designed to investigate contextual support during compound learning. This limits the generalizability of its findings to schema-based learning mainly for two reasons: First, the relationship between the definition and the two words was only established in a non-formalized way, which may have increased the inter-individual variability in the use of this knowledge for the associative encoding of the word pairs. Second, the definition and the sentence condition differ in many more aspects than only the degree of schema-congruency, such as the potential to induce unitized encoding or demands on sentence processing. To overcome these limitations with respect to the current research question, we used an adapted version of the definition-sentence paradigm in which we employed a better operationalization of the semantic relationship between context and compound word, and also established a better control condition (see below).

4.2 The present study

While the majority of the studies on schema-based learning explored how schema knowledge supports memory for single items or for item-context associations, a primary goal of the present study was to assess the mechanisms by which schema knowledge facilitates episodic encoding of two unrelated items into an associative memory representation. Referring to the example in the introduction, how does superordinate semantic knowledge about what “flying”, and “shame” mean (i.e., the schema) benefit the learning of the novel compound word “flight shame”? In the present study, participants were presented with novel German compound words, which were preceded by either a semantically congruent fictional definition, or a neutral definition. Their task was to rate how well the compound word is described by the definition. We assumed that the additional semantic relationship between the fictional definition and the first constituent of the novel compound, i.e., the modifier constituent, in the congruent condition should facilitate the integration of the compound with the knowledge structure provided by the definition, serving as a schema. The congruent definition provided in the current study meets the requirement of the schema definition to structure already known concepts by explaining how both unrelated concepts can be linked to create a new concept, what might influence information processing and learning of the novel compound words. More precisely, the definition provides a template that explains how prior knowledge,

i. e., conceptual knowledge about the compound word constituents, can be used to create a novel concept and how the lexical entries of both word constituents can be linked. This facilitated semantic processing should elevate the activation level of the compound word and by this boost episodic encoding even after a single exposure of the fictional definition. Of note, with this manipulation of schema-congruency, the congruent and the neutral condition only differed in the presence or absence of a semantic relationship between the definition and the modifier constituent of the compound word and all other schema effects were controlled for. As the participants were unaware that a memory test will occur, we could additionally control for the use of intentional encoding strategies that may hinder finding schema-congruency effects on compound learning. Based on prior literature (e.g., Höltje et al., 2019; Staresina et al., 2009), we expected memory performance in the present study to be better for compound words presented together with congruent than neutral definition contexts. Further, the facilitated semantic processing of congruent compounds should be reflected in an attenuated N400 compared to the neutral condition. If the processes indicated by the N400 effect contribute to memory formation, we expect SMEs with a similar temporal and topographical distribution as the N400 effect (N400-SMEs). These N400 SMEs should be modulated by congruency, with larger subsequent memory effects in the congruent condition, as compared to the neutral condition. Conversely, if the processes contributing to successful memory formation are qualitatively distinct from the aforementioned effects, the SME should either temporally precede or follow the N400 effect with a different topographical distribution. One such component is the parietal SME (see Höltje et al., 2019; Kamp et al., 2017). If similar processes support schema-based episodic encoding of novel word associations, as it is the case for single items, we expect a larger, parietally distributed SME for the congruent condition than for the neutral condition. In addition, similar to Kamp et al. (2017), we expect a frontally distributed slow-wave SME for both conditions, reflecting more generally the encoding of associations.

4.3 Methods

4.3.1 Participants

A sample (*N* = 43) of young adults volunteered for this study, having been recruited via flyers and local databases. The required sample size of *N* = 20 was determined with a power analysis (G*Power, Version 3.1.9.4.; Faul et al., 2009) for a one-sided, paired-samples *t* test on the effect-of-interest, i.e., the SME-difference between high-typical congruent and incongruent trials, based on Höltje et al. (2019), $d_z = 0.59$, $\alpha = 0.05$, $1\text{-}\beta = 0.80$. Data from *n* = 13 participants had to be excluded due to failures during recording (*n* = 2), because the stimulus materials were known from another study (*n* = 1), because they reported that they intentionally studied the stimuli or did not give an indication (*n* = 5) or did not provide enough artifact-free trials (*n* = 5). The final sample consisted of *N* = 30 participants (21 females, with an age range from 18 to 31, *Mdn* = 22 years, *SD* = 3.55). All participants performed above chance in the memory test, which was verified with a binomial test (p > .05). All participants were students of Saarland University or volunteers from the community and reported being in good health, not suffering from any neurological or psychiatric conditions and having normal or corrected-to-normal vision. Further, all participants were right-handed, as assessed with the Oldfield Handedness Inventory (Oldfield, 1971), and reported being native speakers of German. Participants gave their informed consent and were reimbursed with 10E/h. Participants were debriefed after the experiment. The experiment was approved by the ethics committee of the Deutsche Gesellschaft für Sprachwissenschaft (#2017-07-180423) and adhered to the Declarations of Helsinki.

4.3.2 Stimulus materials

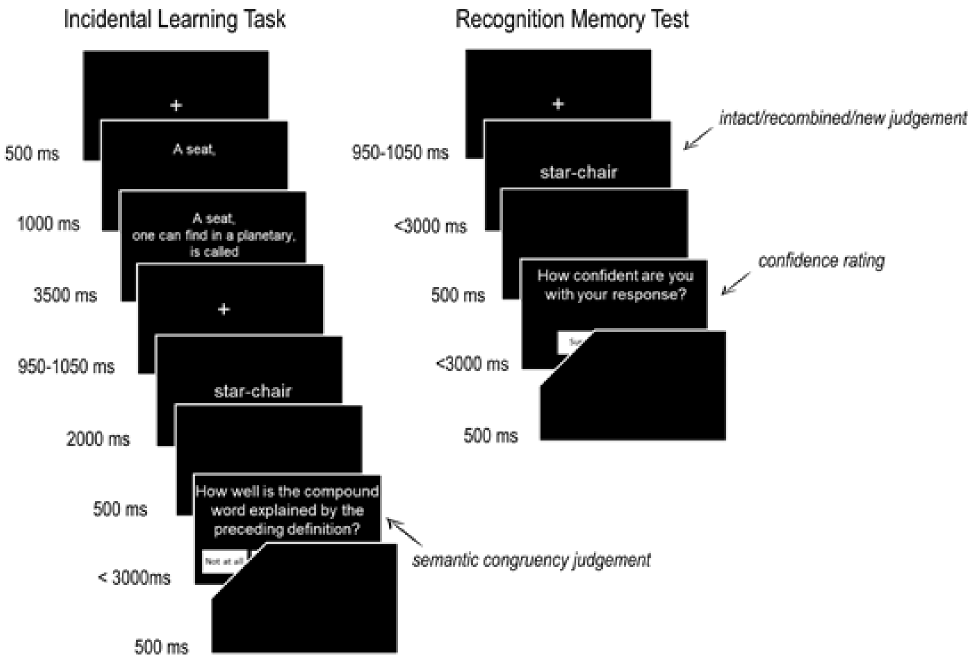
Based on the rating studies presented in chapter 3, we selected 240 compound words out of our stimulus pool, together with their recombined word pair, and 80 new compound words for the final stimulus materials. The selected stimuli were divided into two sets (Set 1 and Set 2), consisting of 120 compound words each. Two encoding lists were created, whereby for the first list, compound words of Set 1 were presented with a congruent context and compound words of Set 2 were presented with a neutral context. This assignment was reversed for the second list. Which encoding list was used varied across participants, whereby both lists were presented approximately equally often. To create lists for the test phase, stimuli were further divided into four subsets of 60 compound words, each, by halving Set 1 and Set 2, respectively. This enabled us to counterbalance the type of presentation of a learned compound word in the test phase, i.e., either as intact or recombined compound word. Consequently, compounds of each subset were presented once as intact and once as recombined across test lists, so that when compound words of Set 1a and Set 2a were presented as intact compound words, the other half (Set 1b and Set 2b) was presented as recombined compound words, and vice versa. Thus, for each encoding list, the same 2 test lists were created, resulting in 4 possible combinations of encoding and test lists. Each test list consisted of 120 intact compound words, 60 recombined compound words, and 80 new (yet unpresented) compound words. The new compound words were identical for each participant. Across all sets (1 and 2) and subsets (1a, 1b, 2a, 2b), there were no statistically reliable differences in normalized lemma frequency of compound constituents, compound word length or semantical relationship between compound word constituents. Further, there were no significant differences in context fit between Set 1 and Set 2 in the encoding lists. Stimulus presentation in the experiment was pseudo-randomized for the encoding and test phase, with the limitation of not more than 3 consecutive trials in the same context condition (encoding phase) or not more than 3 consecutive trials requiring the same response (test phase).

4.3.3 Procedure

After having given their written–informed consent, participants completed several questionnaires, one about their general health, one about demographic aspects and the Oldfield Handedness Inventory (Oldfield, 1971). Next, EEG was applied, and participants were sat in a dimly lighted, sound-absorbing chamber. The experiment was created using E Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). The experiment proper consisted of an incidental encoding phase, a retention interval with a duration of 10 min and a test phase (see Figure 5 for an overview of trial procedures).

Figure 5

Illustration of the trial procedures in the incidental learning task (left) and in the recognition memory test (right)



Note. The depicted example stimuli are English translations of the original German stimulus materials.

All stimuli of encoding and test phase were presented in white font against a black background. During the encoding phase, participants were presented with 240 definitions, half of them congruent and half of them neutral, followed by the respective novel compound word. Participants were instructed to rate on a scale from 1 (not at all) to 4 (absolutely) how well the novel compound word denotes the concept given by the definition. A trial started with a fixation cross, with a duration of 500 ms. Then, the definition was presented stepwise. The noun phrase was presented for 1000 ms, followed by the presentation of the relative clause and the words "is called" for additional 3500 ms. After another fixation cross with a continuously jittered duration from 950 to 1050 ms, the compound word was presented in the center of the screen for 2000 ms. A 500 ms blank screen followed the compound word. Then the answer screen appeared for up to 3 s but was terminated as soon as the participants gave their response. Their task was to rate how well the compound word is described by the preceding definition, providing a measure of the semantic congruency between the definition and the compound word. The answer screen contained the question of how well the compound word is described, as well as the labels for the response scale. Participants responded on a keyboard by using the keys x, c, n, and m with their index and middle fingers of each hand. The scale was ascending for a part of the participants and descending for the other part of participants. A 500 ms blank followed until the next trial started. Before the encoding phase, participants completed 8 practice trials to familiarize with the task.

The encoding phase was followed by a 10-minute retention interval. During this interval, participants performed two distractor tasks. At first, an adapted computerized version of the Digit Symbol Task (Wechsler, 1955) from Häuser et al. (2019) was performed for approximately 5 min, followed by 2.5 min of backwards counting in steps of 3. Only then, participants were told about the upcoming test phase. During the test phase, participants were presented with one of the two test list versions, consisting of 120 intact compound words, 60 recombined compound words and 80 new, i.e., yet unpresented compound words. A trial started with a continuously jittered fixation cross (950–1050 ms). Then, the compound word was presented (for up to 3000 ms), until participants gave their response. Participants gave their answer on a keyboard by using the keys f, j and k to indicate if the compound word was intact, recombined or new. Key assignment was varied by using a latin-square design, ensuring that across participants, each response option was used with an approximately equal frequency. After a 500 ms blank screen, participants were asked to indicate their confidence on the previous response (sure or unsure) using their index fingers, whereby key assignment was ascending for a part of the participants and descending for the other part of participants. The confidence scale remained on the screen for up to 3000 ms or until participants gave their response and was only presented if a response had been logged on the compound word. A trial ended with a blank screen, which was presented for 500 ms. In both, learning and test phase,there were self-paced breaks after 60 trials (encoding phase) or 65 trials (test phase), respectively.

4.3.4 Data acquisition and pre-processing

The EEG was continuously recorded from 28 Ag/AgCl scalp electrodes (Fp1/2, F7/8, F3/4, Fz, FC5/6, FC3/4, FCz, T7/8, C3/4, Cz, CP3/ 4, CPz, P7/8, P3/4, Pz, O1/2, A2) using Brain Vision Recorder 1.0 (Brain Products, Gilching, Germany), whereby all electrodes except from A2 were embedded in an elastic cap (Easycap, Hersching, Germany). Electrode positions followed the extended international 10–20 system (Jasper, 1958). AFz was chosen as ground electrode and two electrodes were applied on the left (A1) and right (A2) mastoid, respectively. Electroocular activity was assessed via four additional electrodes, which were placed above and below the right eye and outside the outer canthi of both eyes. The signal was online referenced to the left mastoid electrode (A1) with the exception of one participant for whom some eye and mastoid electrode channels were interchanged by mistake. For this dataset, data were online referenced to the left canthus electrode. Channel assignment was corrected offline and in an additional step, data were re-referenced to left mastoid. Thus, before the actual preprocessing, all datasets had the same reference. All electrode impedances were kept below 5 kOhm with the exception of the electroocular electrodes' impedances. Data were sampled at 500 Hz. An online filter from 0.016 Hz (time constant 10 s) to 250 Hz was applied. Offline, the data were pre-processed using the EEGLAB (version 2019.1; Delorme & Makeig, 2004) and ERPLAB (version 7.0; Lopez-Calderon & Luck, 2014) toolboxes for MATLAB (MathWorks, Inc.). Therefore, the data from the encoding phase were sampled down to 250 Hz and re-referenced to the average of left and right mastoid. Thereafter, data were filtered for the ICA, using a second-order Butterworth bandpass-filter from 0.5 Hz to 30 Hz (-6dB half-amplitude cutoff, with DC removal). 50 Hz powerline fluctuations were removed with a ParksMcClellan notch filter (default setting order 180; with DC removal; see Parks & McClellan, 1972 for the original algorithm). Data were presegmented by discarding all data points exceeding a time period from 1000 ms before a stimulus onset marker to

2500 ms after a stimulus onset marker. Then, bad segments and experimental breaks, as well as practice trials, were manually discarded. The independent component analysis (ICA) infomax algorithm runica was used to later identify and correct for ocular and muscular artifacts. The resulting IC weights and sphere matrix were then applied to the original data, that were first preprocessed as follows: Data were sampled down to 250 Hz and re-referenced to the average of left and right mastoid. Thereafter, data were filtered with a second-order Butterworth bandpass-filter from 0.05 Hz to 30 Hz (-6dB half-amplitude cutoff, with DC removal). 50 Hz powerline fluctuations were again removed with the Parks-McClellan notch filter. Thereafter, data were presegmented identical to as before the ICA (1000 ms before to 2500 ms after a stimulus onset marker with manual removal of the same bad segments, breaks and practice trials). After, the ICA weights and sphere matrix were applied, and components associated with eye movements and muscular artifacts were identified and removed (up to 5 components per participant). Data were then segmented into epochs of 1996 ms around compound word onset, including a 200 ms baseline. Following baseline correction, a semi-automatic artifact rejection was applied, using the following criteria: a maximally allowed amplitude of -75 up to 75 µV, a maximal difference of values of 100 µV during intervals of 200 ms (window steps of 100 ms),a maximally allowed voltage step of 50µV/s and a maximum of 200 ms of sample points with a deviation from -0.5 to 0.5µV from the maximum voltage in this epoch. To calculate average N400 effects, M = 110.6 trials (SD 15.53, range 56–120) were used in the congruent condition and M = 110.7 trials (SD 15.63, range 60–120) were used in the neutral condition. For the SME analyses, only trials from intact compound words were included and compounds which were used in recombined pairs during the test phase had to be discarded. Hence, average ERPs for the SME analyses were calculated for subsequent hits (intact compound words identified as intact) and subsequent misses (intact compound words identified as recombined or new) for each condition (congruent and neutral context), respectively. Due to an insufficient number of trials per level, the collected confidence ratings were not considered. ERPs of subsequent hits were based on M = 36.33 trials (SD 9.06, range 14–50) in the congruent condition and on M = 26.5 trials (SD 7.83, range 8–41) in the neutral condition. For ERPs of subsequent misses, *M* = 17.4 trials (*SD* 6.96, range 8–38) were used in the congruent condition and *M* = 27.3 trials (*SD* 6.66, range 11–42) were used in the neutral condition. A common phenomenon in studies using SMEs are differential trial numbers for subsequently remembered and forgotten information (e.g., Höltje et al., 2019; Kamp, 2020; Otten & Donchin, 2000) that arise naturally when memory performance is above chance. The differential trial number across conditions results in a worse signal-to-noise ratio for the condition with fewer trials, i.e., the misses, as compared to hits, decreasing statistical power. However, based on a simulation study (Gibney et al., 2020), we assume that our minimal trial criterion is adequate, and we do not face a power issue, given our relatively large sample and the fact that our effects-of-interest, i.e., subsequent memory effects, are rather large effects (e.g., for the parietal SME approx. 2 µV in Kamp et al., 2017 , estimated from Fig. 3A and 1.8 µV for congruent exemplars in Höltje et al., 2019).

4.3.5 Data analysis

For all analyses, the significance criterion of $p < .05$ was applied. Data were analyzed using R (version 3.6.1; R Core Team, 2019) and RStudio (Version 1.2.5001; RStudio Team, 2019) and IBMSPSS statistics (version 26). Whenever non-hypothesis-driven multiple testing was required, the Bonferroni-Holm correction (Holm, 1979) was applied. The reported corrected p-values were calculated with the function p. adjust of the R package stats (R Core Team, 2019). To capture associative memory performance by considering intact and recombined compound words irrespective of correct rejections of new compound words, an associative *Pr* (hits - false alarms) was calculated. Therefore, the associative hit rate was calculated as the amount of compound words, correctly identified as intact, divided by the sum of all intact trials, classified as either intact or recombined. The associative false alarm rate was calculated as the sum of all recombined items classified as intact, divided by the number of recombined items either classified as recombined or intact. Behavioral outliers were defined as extreme values, i.e., with a standardized z-value greater than 3.29 above the mean (Field, 2009, p. 179). No data had to be excluded from behavioral analyses. To analyze ERP data, we pursued a two-step strategy. In a first step (manipulation check), we aimed to evaluate the congruency manipulation. Therefore, we examined N400 congruency effects by comparing ERPs for congruent and neutral trials. As subsequent memory was not relevant for this analysis, all artifact-free trials of the learning phase (subsequently presented as intact or recombined) were used in this analysis. In a second step, our goals were to examine (i) whether the N400 effect was modulated by subsequent memory and (ii) whether we find similar early parietal and late frontal subsequent memory effects as in prior studies (Höltje et al., 2019; Kamp et al., 2017). As subsequent memory effects are difficult to interpret for recombined compound words, due to the different study contexts of the two constituents, only intact trials were used in these subsequent memory analyses. As in previous studies, we used a central-parietal electrode cluster to examine N400 effects (Brothers et al., 2020; Kuperberg et al., 2020), whereby we chose near spatial neighbors in case our electrode montage did not cover the respective positions. The electrode cluster included electrodes Cz, CPz, C3/4. To avoid spatial overlap between the N400SME and the parietal SME, CP3/4-electrodes were omitted from the N400 cluster. For N400-related analyses, we selected an a-priori time window from 300 to 500 ms (e.g., Höltje et al., 2019; Stites et al., 2016). However, due to the slightly delayed N200 preceding the N400, we adjusted this time window post-hoc to 350–500 ms, to obtain a valid measure of the N400 effect (see Yagoubi et al., 2008, for similar adjustments). Interestingly, in the neutral condition, the N400 effect seems to extend until approximately 700 ms, whereby it is attenuated in the congruent condition (see Figure 6). Critically, the topographical distribution of this extended N400 effect resembles the distribution of an N400effect in a remarkable way. Therefore, we analyzed the extended N400 effect in an additional, post-hoc selected time window from 500 to 700 ms (postN400 time window). The same electrode cluster and time windows were used to test whether N400 effects were modulated by subsequent memory (N400SME). Electrodes for the parietal and frontal SME electrode clusters were selected on the basis of Kamp (2020), whereby we again chose near spatial neighbors in case our electrode montage did not cover the respective positions. The parietal SME cluster included electrodes CP3/ 4, P7/8, P3/4, Pz, O1/2. Consistent with prior research on the parietal SME and the N400 (Höltje et al., 2019; Packard et al., 2017), the same time window as for the a priori defined N400 and the post-N400 time window were chosen for the analyses of the parietal SME. However, similar to the N400 effect, visual inspection revealed that the maximum of the parietal SME was shifted in time to 700 – 900 ms. The temporal shifts of the N400 and SME effects were presumably caused by the more multifaceted congruency manipulation, requiring prolonged semantic processing, as compared to the Höltje et al. (2019) study in which merely short category cues were used as context manipulations. Therefore, we additionally analyzed the parietal SME in the post-hoc defined time window from 700 to 900 ms. The frontal SME was analyzed on a frontal-central cluster, including electrodes Fp1/2, F7/8, F3/4, Fz, FC5/6, FC3/4, FCz, in a time window from 900 to 1200 ms (Höltje et al., 2019). Note that we re-ran the analyses for a time window from 1200 to 1796 ms, similar to Kamp et al. (2017), yielding qualitatively identical results, which are not reported here. All topographical profile analyses were conducted with the electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. As sphericity is usually violated in EEG data, we used the multivariate approach of repeated measure analysis of variance (MANOVA), which is more robust against such violations of sphericity (Dien & Santuzzi, 2005; Picton et al., 2000). For the sake of readability, we only report significant effects including the factors congruency or memory. Significant effects are further explored in follow-up MANOVAs and paired-samples *t* tests. As measures of effect size, we report Hedges' g_{av} for effects from paired samples t tests with the formula provided in the spreadsheet (Version 5; Lakens, 2013) and Pillai's trace, which is identical to partial eta squared (η2), for multivariate analyses of variance (MANOVAs), respectively.

4.4 Results

4.4.1 Behavioral results for the encoding phase

We compared the responses for the rating during the encoding phase between the congruent and the neutral condition with a paired-samples *t* test. As expected, participants rated compound words in the congruent condition as being explained better than compound words in the neutral condition, $t(29) = 24.10$, $p < .001$, $g_{av} = 4.82$ ($M_{congruent} = 3.20$, $SD = 0.34$; $M_{neutral} = 1.68$, $SD = 0.26$).

4.4.2 Behavioral results for the test phase

To test if and how congruency modulates memory performance, we calculated a Congruency (congruent, neutral) × Type (hit, false alarm) MANOVA. This analysis revealed a main effect of congruency, $Pillai = 0.57$, $F(1, 29) = 38.67$, $p < .001$, a main effect of type, $Pillai = 0.86$, $F(1, 29) = 173.94$, $p < .001$, and a significant interaction, $Pillai = 0.28$, $F(1, 29) = 11.16$, $p = .002$. Further examination revealed that only hit rates differed significantly across congruency conditions, $t(29) = 10.64$, $p < .001$, $g_{av} = 1.22$ (one-sided, $M_{congruent} = 0.76$, $SD = 0.12$, $Mneutral = 0.60$, $SD = 0.13$), whereas no such differences were found for false alarm rates, $t(29) = 1.50$, $p = .144$, $g_{av} = 0.24$ (two-sided, $M_{congruent} = 0.36$, $SD = 0.19$, $M_{neutral} = 0.31$, $SD = 0.16$). Consequently, associative memory performance, indicated by association-based *Pr*, was higher in the congruent than in the neutral condition, $t(29) = 3.34$, $p = .001$, $g_{av} = 0.64$ (one-sided, $M_{congruent} = 0.40$, $SD = 0.19$, $M_{neutral} = 0.29$, $SD = 0.15$).

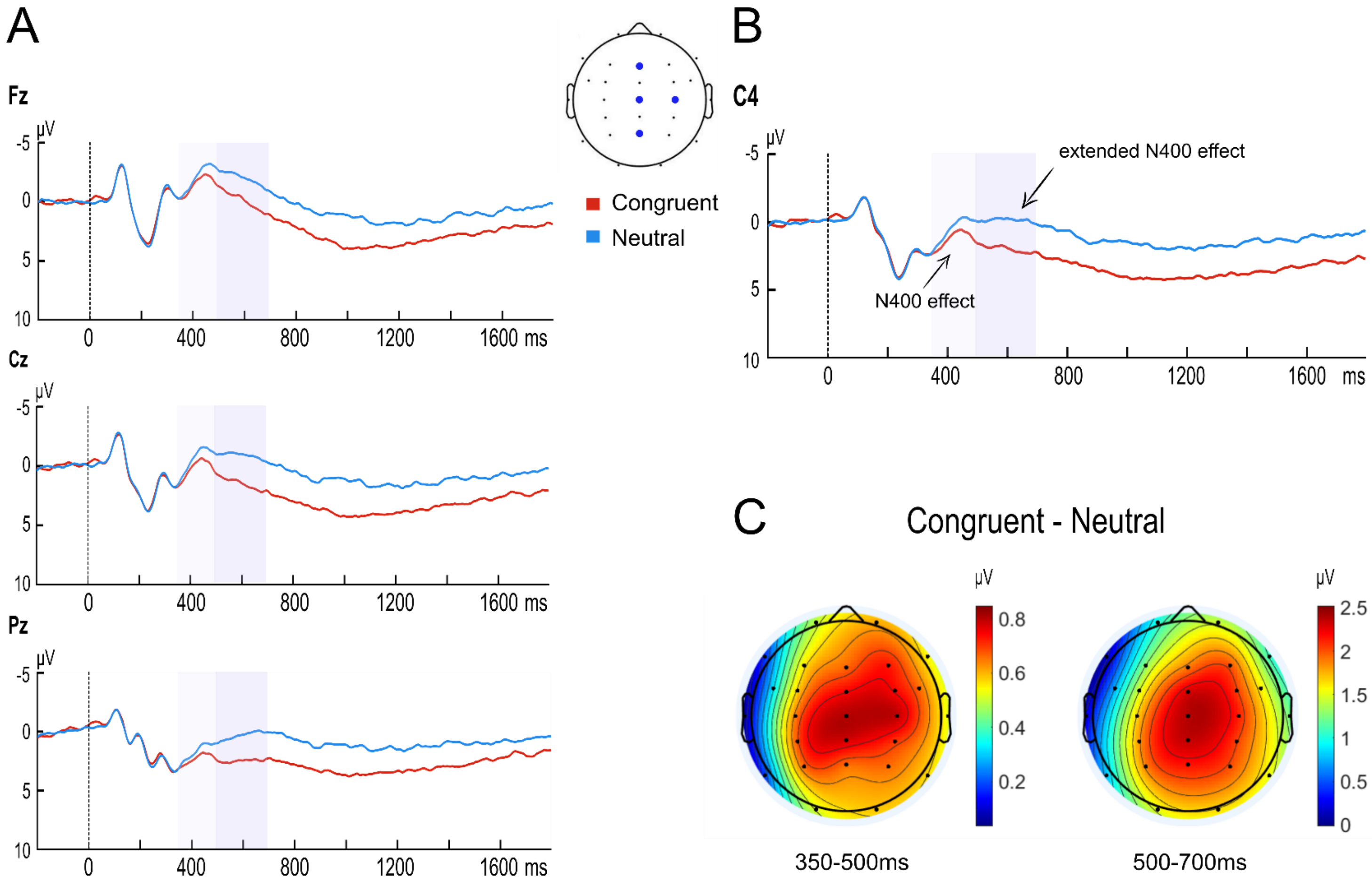
4.4.3 ERP results

Figure 6 depicts the grand average ERP waveforms elicited by the compound words in the encoding phase. Effects of congruency for the N400 start to emerge at approximately 350 ms with a right-central topography. The N400 is attenuated in the congruent condition relative to the neutral condition. Interestingly, in the neutral condition, the N400 effect appears to be extended until approximately 700 ms, whereby it is attenuated in the congruent condition. The subsequent positivity seems to start earlier in the congruent condition. The subsequent memory effects in both conditions are illustrated in Figure 7. Subsequently remembered compound words show more positive ERP deflections than subsequently forgotten ones. In the congruent context condition, ERPs to subsequent hits and misses start to diverge at approximately 300 ms at parietal sites, with maximal effects occurring between 500 and 900 ms post-stimulus. In addition, there is a late SME, which is present at frontal recording sites in both conditions. The late frontal effect appears to be largest in a time window between 900 and Figure 7. Subsequent Memory Effects. Panel A depicts subsequent memory effects in the congruent and in the neutral condition. The topographic maps of the sub sequent memory effect (hit - misses) in all four analyzed time windows are illustrated in Panel B, for each condition, separately.

These observations were examined in a series of statistical analyses, whereby we first present a-priori defined analyses, followed by post-hoc analyses (see Figure 8 for an overview of the analysis approach).

Figure 6

N400 Effects



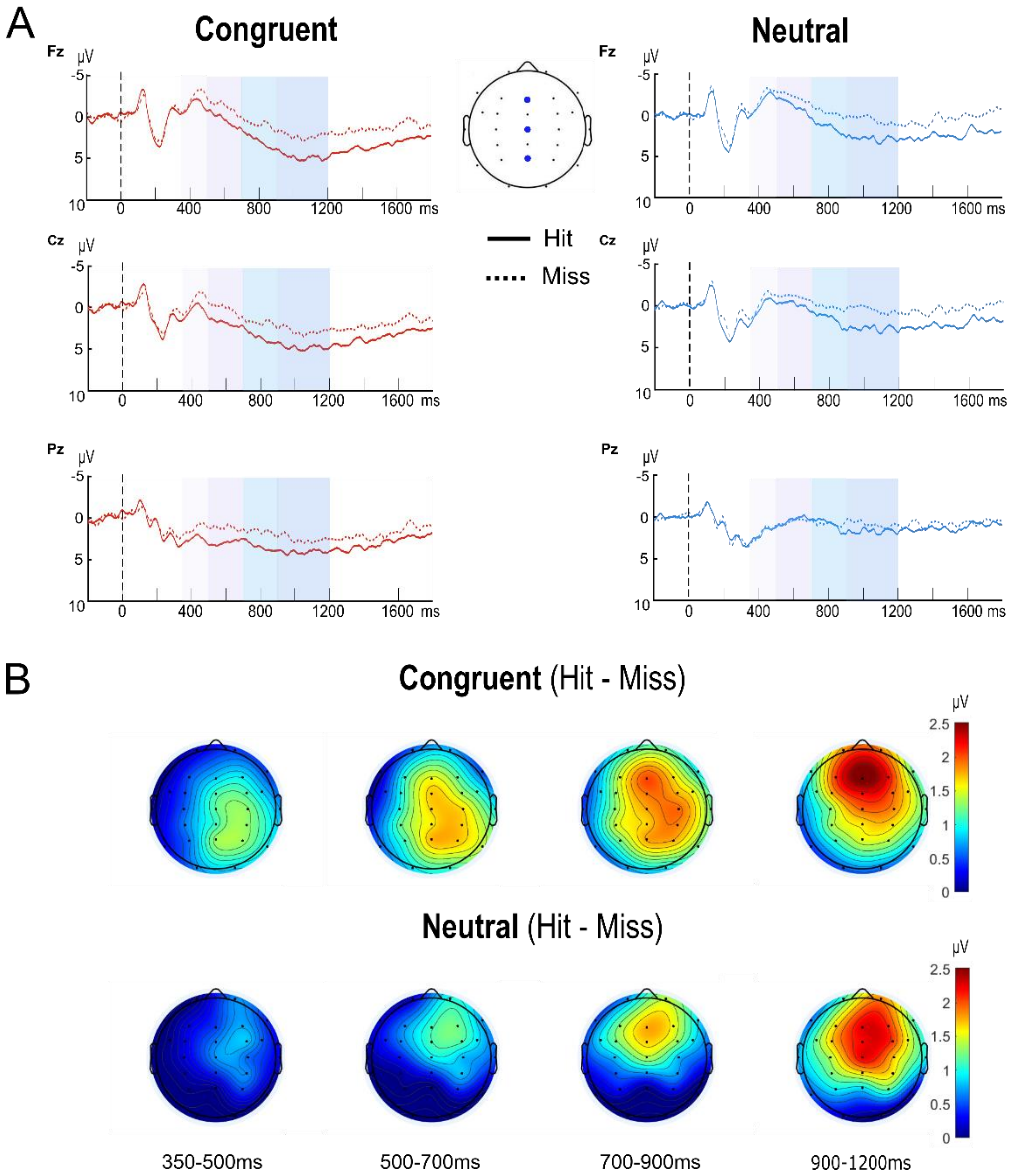
Note. Panel A depicts ERP waveforms of the encoding phase for compound words presented in a congruent and in a neutral context. Panel B depicts N400 effects at electrode C4, where the N400 and the extended N400 effect can be observed particularly well. Panel C depicts the topography of the N400 effect (congruent - neutral) in both time windows.

4.4.3.1 N400 congruency and subsequent memory effects (350–500 ms)

A paired-samples *t* test on the central-parietal N400-cluster in the apriori defined time window from 350 to 500 ms was conducted to compare mean amplitudes in the congruent and in the neutral condition. Consistent with our hypotheses, mean amplitudes were more positive in the congruent condition, as compared to the neutral condition, $t(29) = 3.21$, $p = .002$ (one-sided), $g_{av} = 0.20$ (see Figure 9). In a next analysis step, we aimed to investigate whether the N400 contributes to successful memory formation. Therefore, a Memory × Congruency-MANOVA was calculated in the a-priori defined N400 time window from 350 to 500 ms on the central-parietal N400 electrode cluster. This analysis only revealed a significant main effect of memory, *Pillai* = 0.18, $F(1, 29) = 6.18$, $p = .019$, with in general more positive amplitudes for remembered, as compared to forgotten compound words, irrespective of congruency. Thus, we observed the predicted N400 effect with less negative amplitudes in the congruent, as compared to the neutral condition, but the N400-SME was not modulated by congruency (Figure 9).

Figure 7

Subsequent memory effects.



Note. Panel A depicts subsequent memory effects in the congruent and in the neutral condition. The topographic maps of the sub sequent memory effect (hit - misses) in all four analyzed time windows are illustrated in Panel B, for each condition, separately.

4.4.3.2 Parietal subsequent memory effects in the N400 (350–500 ms) and post-N400 (500–700 ms) time windows

A Memory × Congruency × Time Window-MANOVA on the parietal electrode cluster revealed a significant main effect of time window, *Pillai* = 0.26, $F(1, 29) = 10.00$, $p = .004$, a significant main effect of congruency, *Pillai* = 0.23, $F(1, 29) = 8.66$, $p = .006$, a significant main effect of memory, *Pillai* = 0.22, $F(1, 29) = 8.18$, $p = .008$, as well as a significant interaction of time window and congruency, *Pillai* = 0.44, $F(1, 29) = 22.96$, $p < .001$, and a significant interaction of congruency and memory, *Pillai* = 0.13, $F(1, 29) = 4.35$, $p = .046$. To resolve the significant interaction of time window and congruency, data were averaged over the factor memory to calculate Bonferroni-Holm corrected, follow-up paired-samples *t* tests on congruency for each time window, separately. Those revealed more positive amplitudes in the congruent than in the neutral condition in the later time window, $t(29) = 4.05$, $p < .001$, $g_{av} = 0.48$, but not in the earlier time window, $t(29) = 0.25$, $p = .807$, $g_{av} = 0.02$. To resolve the significant interaction of congruency and memory, data from both time windows were averaged and follow-up paired-samples *t* tests on memory were calculated for the congruent and the neutral condition, separately. Hereby, a significant parietal subsequent memory effect with more positive amplitudes for remembered, as compared to forgotten trials was found in the congruent condition, $t(29) = 3.23$, $p = .002$ (one-sided), $g_{av} = 0.37$, but not in the neutral condition, $t(29) = 0.25$, $p = .808$ (two-sided), $g_{av} = 0.03$ (see Figure 9). Thus, consistent with our hypotheses, we found a parietal SME that was modulated by congruency being only statistically reliable in the congruent condition.

4.4.3.3 Frontal subsequent memory effects (900–1200 ms)

A Memory × Congruency-MANOVA on frontal electrodes revealed a significant main effect of congruency, *Pillai* = 0.28, $F(1, 29) = 11.29$, $p = .002$, with more positive amplitudes in the congruent, as compared to the neutral condition, and a significant main effect of memory, *Pillai* = 0.49, $F(1, 29) = 27.58$, $p < .001$, with more positive amplitudes for remembered than for forgotten trials. Consequently, consistent with our hypotheses, we found a late frontal SME that is independent of congruency (see Figure 9).

4.4.3.4 Post-hoc analyses of the extended N400 effect (post-N400 time interval)

An additional paired-samples *t* test on the central-parietal N400 cluster was conducted in the post-N400 time window (from 500 to 700 ms), using the same central-parietal electrode cluster. Again, mean amplitudes were more positive in the congruent condition, as compared to the neutral condition, $t(29) = 6.31$, $p < .001$ (two-sided), $g_{av} = 0.54$ (see Figure 6). To explore whether this effect was modulated by subsequent memory, we calculated a Memory × Congruency-MANOVA. As in the N400 time interval, this analysis revealed a significant main effect of congruency, *Pillai* = 0.27, $F(1, 29) = 10.52$, $p = .003$, as well as a significant main effect of memory, *Pillai* = 0.32, $F(1, 29) = 13.70$, $p = .001$ (see Fig. 5). Thus, as in the N400 time interval, we found congruency effects and subsequent memory effect also in the post-N400 time window, with no interaction between the two factors.

4.4.3.5 Post-hoc analyses of the late parietal subsequent memory effect

An additional post-hoc analysis on the parietal cluster in the time window from 700 to 900 ms was computed. A Memory × Congruency MANOVA on the parietal electrode cluster revealed a significant main effect of congruency, *Pillai* = 0.46, $F(1, 29) = 25.00$, $p < .001$, a significant main effect of memory, *Pillai* = 0.17, $F(1, 29) = 5.98$, $p = .021$, and a significant interaction of congruency and memory, *Pillai* = 0.16, $F(1, 29) = 5.34$, $p = .028$. To resolve the significant interaction, Bonferroni-Holm-corrected, follow-up paired-samples *t* tests on subsequent memory were calculated for the congruent and the neutral condition, separately. This analysis revealed a significant parietal subsequent memory effect in the congruent condition, $t(29) = 3.00$, $p = .010$ (two-sided), $g_{av} = 0.44$, but not in the neutral condition, $t(29) = 0.24$, $p = .811$ (two-sided), $g_{av} = 0.03$ (see Fig. 5). Thus, comparable to the earlier parietal SME (350–700 ms), the parietal SME in this later time window was modulated by congruency.

Figure 8

Overview of the ERP analysis strategy and the results of the statistical analyses.

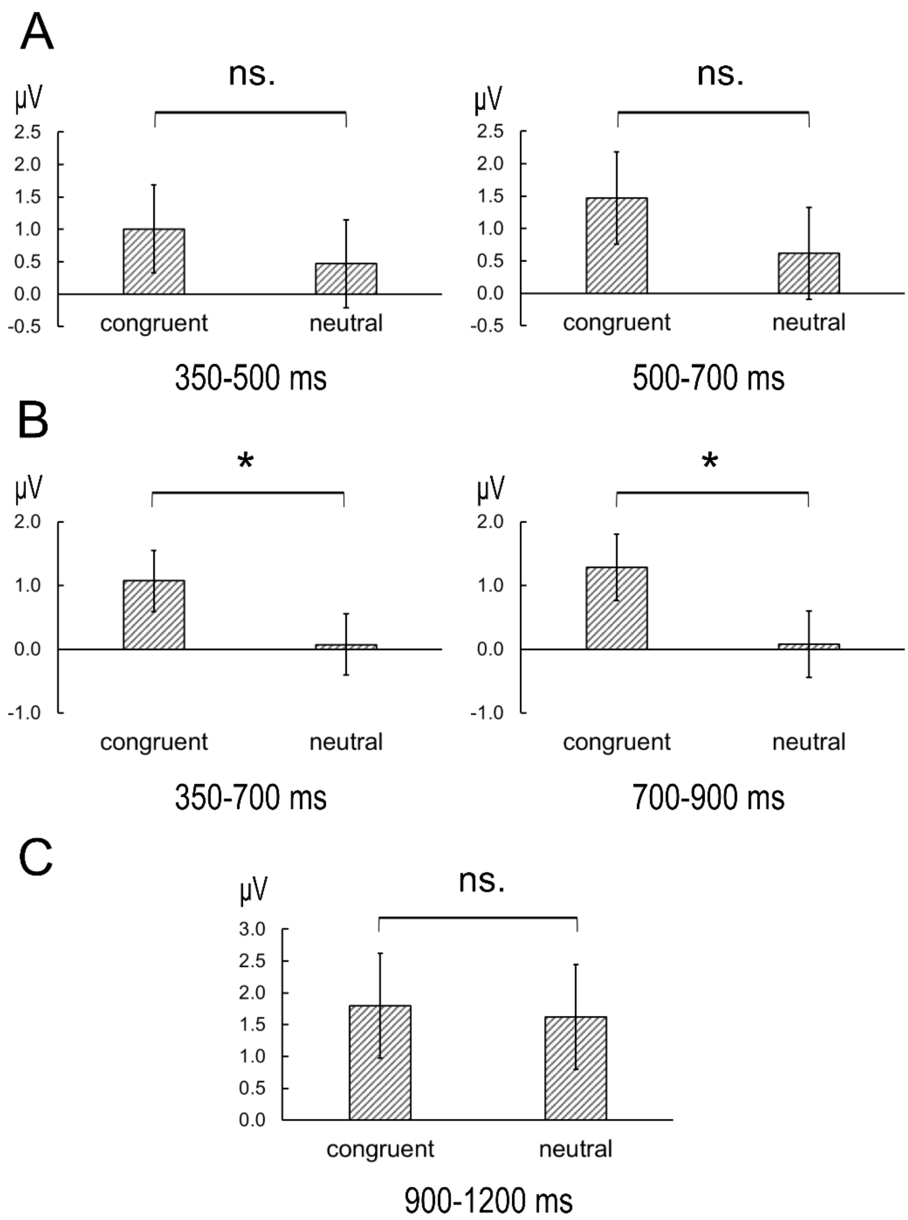
Type of effect	350-500ms	500-700ms	700-900ms	900-1200ms
Congruency (All trials)	**A-priori: N400 effect** Less negative amplitudes in congruent vs. neutral context	**Post-hoc: N400 effect** Less negative amplitudes in congruent vs. neutral context		
Subsequent memory and congruency (Subsequently intact trials)	**A-priori: N400-SME** More positive amplitudes for subsequently remembered vs. forgotten compound words, no congruency modulation of effect	**Post-hoc: N400-SME** More positive amplitudes for subsequently remembered vs. forgotten compound words, no congruency modulation of effect	**Post-hoc: parietal SME** More positive amplitudes for subsequently remembered vs. forgotten compound words in congruent condition	**A-priori: frontal SME** More positive amplitudes for subsequently remembered vs. forgotten compound words, no congruency modulation of effect
	A-priori: parietal SME More positive amplitudes for subsequently remembered vs. forgotten compound words in congruent condition (spans 350-700ms)			

4.4.3.6 *Topographical profile analyses*

To explore whether the topographic profiles of N400 congruency effects and subsequent memory effects differ qualitatively, and not merely in relative strength (Urbach & Kutas, 2002; 2006), we vector-scaled the data according to McCarthy and Wood (1985) and calculated Effect × Anteriority × Laterality-MANOVAs for the respective effects and time windows. We report only significant effects including the effect factor. Here, we were particularly interested in three topographical contrasts: (A) the contrast between the condition unspecific N400SME and the N400 congruency effect (in the N400 and post-N400 time interval), (B) the contrast between the parietal SME in the congruent condition and the N400 congruency effect (in the N400 and post-N400 time interval) (C) the frontal SME (collapsed across both levels of congruency) and the parietal SME. To calculate congruency effects, the mean amplitude difference between congruent and neutral trials was computed for each participant and time window, separately. To calculate SMEs, mean amplitudes for forgotten trials were subtracted from mean amplitudes for remembered trials. Thereafter, all scores were vector-scaled. In the following analyses, only significant effects including the effects factor are reported. A: With this first contrast, we examined whether the spatial distribution of the N400-SME is comparable to the spatial distribution of the N400 congruency effect, which, together with their similar temporal characteristics, would support the assumption that the N400-SME is indeed functionally equivalent with the N400 effect. MANOVAs with the factors effect (N400-SME, N400 congruency effect), antPos (anterior, central, posterior) and laterality (left, mid, right) did not reveal significant three-way interactions, neither in the N400, *Pillai* = 0.08, $F(4, 26) = 0.55$, $p = .698$, nor in the post-N400 time interval, *Pillai* = 0.17, $F(4, 26) = 1.30$, $p = .297$. No other interaction including the effects factor approached significance. Thus, there is no evidence for different topographic distributions of the N400-SME and the N400-congruency effect, and thus no empirical evidence for functionally independent ERPs in both time intervals. B: Next, we aimed to investigate whether the parietal SME and the N400 congruency effect differ in their topographic profiles, what, together with their different temporal expansion, would support the assumption that the parietal SME and the N400 congruency effect differ functionally. For the parietal SME in the congruent condition, we chose the time window from 700 to 900 ms, because the effect was largest in this time window. MANOVAs with the factors effect (congruent parietal SME, N400 congruency effect), antPos (anterior, central, posterior) and laterality (left, mid, right) revealed a marginal significant three-way interaction, *Pillai* = 0.30, $F(4, 26) = 2.72$, $p = .051$ in the N400 time interval and a significant three-way interaction in the post-N400 time window, *Pillai* = 0.44, $F(4, 26) = 5.20$, $p = .003$. Thus, the parietal SME in the congruent condition differs qualitatively in its topographic profile from the N400 congruency effect in both N400 time intervals. C: Lastly, we aimed to test if the parietal SME in the congruent condition (700–900 ms) and the late frontal SME (900–1200 ms, collapsed across both levels of congruency) differ in their topographic profiles, providing evidence against the parietal SME being a mere continuation of the frontal SME. An Effect (congruent parietal SME, late frontal SME) × AntPos (anterior, central, posterior) × Laterality (left, mid, right)-MANOVA revealed a significant three-way interaction *Pillai* = 0.32, $F(4, 26) = 3.12$, $p = .032$, suggesting that the topographic profiles of the parietal and frontal SMEs differ, as well. To summarize: While we did not find topographical differences between the N400-SME and the N400-congruency effect as well as between the extended N400 SME and the extended N400 congruency effect, the parietal SME in the congruent condition could be topographically differentiated from the N400-congruency effect, the extended N400 congruency effect and the late frontal SME.

Figure 9

ERP Amplitudes of Subsequent Memory Effects.



Note. Panel A shows hit minus miss-mean amplitude measures for the N400 cluster in the earlier (350–500 ms) and the later (500–700 ms) N400 time window in each condition. Hit minus miss-mean amplitude measures of the parietal SME in the earlier (350–700 ms) and the later (700–900 ms) time window in the parietal cluster are illustrated in Panel B, for each condition, separately. Panel C shows the hit minus miss-mean amplitude measures in each condition for the frontal cluster from 900 to 1200 ms. The asterisk marks statistically significant effects ($p < .05$). Error bars represent +/- 1 standard error of the mean difference for the across-conditions comparison in all diagrams.

4.5 Discussion

An extensive number of studies have demonstrated that events that are congruent with a given schema are remembered better than incongruent events. However, the mechanisms by which prior semantic knowledge facilitates episodic encoding of new information still need to be specified. In the present study, we extend the schema framework to the learning of novel word associations, i.e., compound words, and explored whether schema knowledge supports the encoding of two previously unrelated words. We manipulated the semantic relationship between novel compound words and a fictional definition, from which we assume that it fulfills the requirements of a schema. Thus, we explored whether and how a strong semantic relationship between the schema context and the compound word constituents contributes to episodic memory formation.

4.5.1 Behavioral results

The finding that events that are congruent with a given schema are remembered better is well established in the neuropsychological literature, and this congruency effect has been reported for a wide range of tasks and modalities (e.g., Bein et al., 2014; Pichert & Anderson, 1977; Staresina et al., 2009; van Kesteren et al., 2013). Whilst most of these studies focus on the learning of single items, several studies reported beneficial effects of schema-congruency for the learning of associations (e.g., Bein et al., 2014; Staresina et al., 2009; van Kesteren et al., 2013). The current study focusses on the learning of novel word associations, i.e., novel compound words. In this setting, the constituting items are already known and integrated into prior knowledge structures whilst a novel association between these items must be acquired. Further, the current study embeds the semantic congruency manipulation in a rich linguistic context, which we argue is a more ecologically valid operationalization of semantic knowledge use as e.g., word-color associations (Staresina et al., 2009). The memory advantage for schema-congruent events in the congruent condition, which we found in the present study, is well in line with the putative easier integration of information that matches representations in semantic memory. This might in turn lead to richer and more elaborated memory traces, which are better accessible in a subsequent memory test (Craik & Tulving, 1975). To ensure that participants could not solve the memory task by relying on item information only, i.e., by assessing the memory strength of the compound constituents, but were enforced to remember the exact combination of the word constituents, recombined compound words were presented during the test phase (together with not yet presented compound words). The finding that between-condition differences were larger for hits than for false alarms to recombined pairs suggests that the congruent context did not just induce a bias to endorse already presented single words as “old” by means of item memory, but rather boosts episodic encoding of the underlying association. How exactly schema congruency fosters the creation of an associative memory representation, e.g., by creating a semantic link between underlying concepts (Boutonnet et al., 2014) cannot be determined based on data from the current study.

4.5.2 The N400 and the extended N400 congruency effect

In the current study, we found the expected N400 effect from 350 to 500 ms post-stimulus, i.e., an attenuation of the N400 in the congruent condition relative to the neutral condition. This effect is consistent with a large number of studies showing similar semantic congruency effects for the N400(Bridger et al., 2012; DeLong et al., 2005; Kutas & Hillyard, 1980; Van Petten & Luka, 2012; see Kutas & Federmeier, 2011 for a review). It could be argued that the two context conditions did not only differ in the semantic congruency between the definition and the compound word, but also in the amount of semantic content of the definitions themselves. Congruent contexts may have been richer in content and thus may have allowed to better predict the target words (see Federmeier et al., 2007, as an example) and these differences may have facilitated semantic processing and boosted episodic encoding of the compound words. To test whether both context types differed in their predictability, we conducted an additional rating study (see chapter 3.2.4) in which we presented congruent and neutral definitions without the compound word and asked participants to indicate how well they could imagine something from the definition. The rationale behind this approach was to check if the congruent condition induces more constraint and thus predictive potential than the neutral condition, when presented without the compound word. As a cloze study is not suitable to estimate constraint for novel compound words, constraint was operationalized as imageability, i.e., how well someone could imagine something from the definition. As there was no significant difference between conditions for these ratings, $t(239) = 1.90$, $p = .058$, $g_{av} = 0.15$ ($M_{congruent} = 2.97$, $SD = 0.47$, $M_{neutral} = 3.04$, $SD = 0.45$), with an opposite numerical trend, we conclude that differences in the predictability of the compound words between conditions (as operationalized in this rating study) cannot account for the N400 effects. As this alternative explanation can be ruled out based on these data, we feel safe to interpret the N400 as a result of facilitated semantic processing of the compound words, due to the preceding definition. Of note, an N400-congruency effect (350–500 ms) was not found when subsequent memory was considered in the analysis. Critically, subsequent memory ERPs only included trials of compound words that were presented identically during the test phase (intact compound words), automatically halving the number of potentially to-be-analyzed trials. This reduction in the signal-to-noise ratio might have prevented the detection of the effect. However, as the N400 congruency effect in the current study is in line with a plethora of studies on semantic priming and the N400 (e.g., Boutonnet et al., 2014; Holcomb, 1993), we deem it as a reliable measure of semantic congruency. Visual inspection of the waveforms revealed an additional, extended N400 congruency effect following the N400. A post-hoc analysis on this effect (500–700 ms) revealed more positive amplitudes in the congruent, as compared to the neutral condition, with similar polarity and distribution as the N400. Although it is not possible to identify this effect as an additional N400 effect based on the data at hand, it is tempting to speculate that the temporal extension of the N400 could be the result of the combinatorial processing of the compound word constituents that is required to compute a whole word meaning. Dual-route approaches of compound word processing (Isel et al., 2003; Koester et al., 2007; Koester et al., 2004; Sandra, 1990; Zwitserlood, 1994; Libben, 2006) assume that already established compound words may be represented in a single lexical entry or are decomposed and analyzed as individual constituents via combinatorial mechanisms. These processes occur in a parallel fashion (e.g., Caramazza et al., 1988). As the compound words in the present study are novel, there cannot yet exist an accessible lexical entry (Libben, 2006). Accordingly, we assume that all novel compound words in the current study must be decomposed, and the conceptual representations of its constituents must be accessed in order to be integrated to a whole word meaning (see Gagné & Spalding, 2009). Interestingly, there is evidence that the N400 is sensitive to this form of lexical-semantic integration of compound word constituents (Koester et al., 2007; 2009). Thus, the cumbersome semantic processing of the novel compound words, i.e., the retrieval of conceptual information of the constituents from long-term memory and its semantic integration, might provide an explanation for an extended N400 effect in the current study. Unfortunately, to the best of our knowledge, there is no study directly investigating the temporal characteristics of processing of novel compound words with ERPs in the visual domain. Thus, this topic should be addressed in future studies.

4.5.3 Subsequent memory modulations of the N400 and the extended N400 congruency effect

Interestingly, in the present study, we found evidence that the semantic facilitation, reflected by the N400 effect, contributes to successful memory formation, as there was an N400-SME in both context conditions, which did not differ from the N400 congruency effect in its scalp topography. A similar N400-SME was obtained for a schema congruency manipulation in Neville et al. (1986), although in this latter study, the N400-SME was modulated by congruency (i.e., larger in the congruent condition). However, as this analysis was based on data from a very small sample (n = 5), these results should be interpreted with caution. Critically, as the N400-SME in the current study did not vary across conditions, it cannot account for the behavioral memory advantage in the congruent condition. This complicates its functional interpretation at first glance. However, as we already argued above, the N400 has also been found to be sensitive to the ease of semantic integration of compound word constituents (Koester et al., 2007; 2009). Semantic integration might have been facilitated by priming effects of the head noun in both contexts and the modifier noun in the congruent context. The ease of semantic integration of the constituents might benefit memory formation and this would then be reflected in the N400-SME. The attentive reader might wonder why the N400-SME is then not larger in the congruent condition, where there is the additional modifier priming effect next to the context-independent head priming effect. Theories on conceptual combination of modifier-head phrases assume that the modifier is used to retrieve information about which type of thematic relationship is frequently used when this word is used as a modifier, whereby the head is used to select from competing relations (Gagné, 2002). Thus, the data of the current study might suggest that only head priming is reflected in the N400-SME, probably indicating the facilitated selection of a fitting relation, whereby it remains unclear why the additional modifier priming in the congruent condition is not reflected in larger N400-SMEs. However, semantic integration of the novel compound words might have been influenced by other factors. The Competition Among Relations in Nominals (CARIN) theory (Gagné & Shoben, 1997; Gagné, 2002) assumes that to compute the meaning of a modifier-noun phrase, concepts are combined by selecting an adequate relation linking both concepts. Here, several possible relations can be distinguished, as e.g., made of: snowball is a ball made of snow (Gagné & Spalding, 2009). Which relations come into consideration when a combination is encountered is determined by the modifier. It is assumed that the modifier contains information about the frequency with which it is used within a particular relation in already known conceptual combinations, i.e., a relational distribution. A modifier’s relational distribution influences how easy a combination is interpreted, whereby high-frequency relations are easier to interpret than low-frequency relations (Gagné, 2002). Whereas the modifier determines which relations are considered, the head noun is used to validate competing relations (Gagné, 2002). Critically, albeit relation availability is influenced by the predictability of an additionally presented linguistic context, pre-existing differences in relation availability are not overridden by the context (Gagné & Spalding, 2004). Following Middleton et al. (2011), this approach is referred to as the generation hypothesis, i.e., “that the initial interpretation of novel combinations in context is based on the generation of a meaning” (p. 809), which is influenced by modifier relation frequency. An alternative approach, the anaphor resolution hypothesis, assumes that when a context is present, it is first attempted to link the combination to a referent in the context and sense generation, in terms of the generation hypothesis, is only engaged if no referent is provided by the discourse (Middleton et al., 2011, cf. Gerrig & Bortfeld, 1999). However, Middleton et al. (2011) provide empirical evidence in favor of a third account, the dual-process hypothesis, that assumes that when a compound word is encountered with a context, sense generation, based on the constituents, and context-driven anaphor resolution, i.e., linking the combination to an earlier discourse referent provided by a context, run in parallel and “either or both may inform the initial interpretation of a novel combination in context” (p. 809). Consistent with the dual-process hypothesis, the N400-SME might reflect memory relevant but context-independent differences in the ease of semantic integration of the compound words via sense generation, i.e., the availability of modifier relation information and morpho-semantic knowledge about the head noun, required to select an adequate relation. The fact that the same compound words with their modifier relation frequency distributions were used in both context conditions might explain the context-independency of the N400-SME. In contrast, the simultaneously onsetting and long-lasting parietal SME probably reflects context-dependent processes as involved in context-driven anaphor resolution, i.e., linking the combination to an earlier discourse referent, provided by the context, in the service of memory formation. This fits well with the context-dependency of the parietal SME.

4.5.4 The early parietal subsequent memory effect

The parietal SME was larger for compound words that were preceded by a congruent, compared to a neutral context. This effect emerged at approximately 350 ms and reached largest amplitudes in the time window from 700 to 900 ms. Notably, the N400 effect, as well as the extended N400 effect and the parietal SME in the 700 to 900 ms time interval showed qualitatively distinct scalp topographies, which suggests that both effects can be functionally dissociated. This finding resembles the results of the Kamp et al. (2017) study, which explored the learning of associations using a definition-sentence paradigm (e.g., Bader et al., 2010; Quamme et al., 2007; Wiegand et al., 2010). As compared to this study, in which the relationship between a congruent context and the compound word was only broadly defined (Kamp et al., 2017), we carefully manipulated schema congruency as the semantic relationship (verified in a rating study) between a fictional definition and the modifier constituent of the novel compound word (congruent context condition) and included a neatly matched neutral control condition. As a similar SME was found also only for words which were congruent with a preceding category cue in a recent study on schema-based learning (Höltje et al., 2019), the parietal SME might reflect some form of integration of semantic information with congruent events in the service of successful memory encoding. The present results confirm and extend these findings in showing that schema-based learning boosts not only learning of single items, but also learning of associations by means of similar mechanisms. We argue that there are at least two different processing mechanisms, which might account for the parietal SME. Subsequent memory effects with similar temporal and spatial characteristics have been reported in memory tasks probing memory for single items or item-specific details (Kamp et al., 2017; Karis et al., 1984) or memory for stimuli which are distinctive in their processing context (Fabiani & Donchin, 1995), like emotionally negative items (Kamp et al., 2015) or pictorial stimuli which are retrieved based on verbal probes (Gonsalves & Paller, 2000). These parietal SMEs have been originally described as modulations of the P300 (Sutton et al., 1965). It has been shown that low probability events elicit a P300 and, probably due to their distinctiveness, are later better remembered (Fabiani et al., 1986; Karis et al., 1984), i.e., the von Restorff effect (von Restorff, 1933). However, this distinctiveness explanation of the parietal SME is more plausible when isolated events must be processed, which is not the case in the present study. However, the congruent condition provides a better framework to integrate the meaning of both compound word constituents into a joint, single item representation than the neutral condition. Thus, it is possible that the selective parietal SME in the congruent condition reflects the item-specific processing of the novel compound word, resulting in a single item representation. Alternatively, it is also conceivable that the parietal SME reflects processes normally indicative for the P600. The P600 is a positive ERP component with a centro-parietal distribution that onsets between 500 and

1000 ms after the onset of the critical word and has originally been related to syntactic processing during language comprehension (Friederici et al., 2002; Hagoort et al., 1993; Osterhout & Holcomb, 1992). This functional interpretation of the P600 has been refined recently, as the P600 has also been observed for forms of semantic processing, as in joke comprehension (Coulson & Kutas, 2001), irony (Regel et al., 2011) and the processing of metaphors (Bambini et al., 2016). In their Retrieval Integration (RI) account, Brouwer et al. (2012) assume that the P600 reflects the "construction, revision, or updating of a mental representation of what is being communicated" (Brouwer et al., 2012, p.137). Thus, the P600 might reflect prolonged attempts to make sense of an input that initially produced a conflict (Kuperberg et al., 2020). We argue that variations of the P600 could account for the parietal SME in the congruent condition, as similar to metaphors; the literal meaning of the compound word must be overridden in favor of the whole word meaning provided by the context. In doing so, the novel whole-word concept is created and mapped onto the word form of the novel compound word, updating its mental representation. These processes are only initiated in the congruent condition, where it is possible to integrate the conceptual combination into prior knowledge structures by creating a conceptual compound representation, which could be beneficial for memory formation, and which is not the case in the neutral condition. The more positive waveforms for subsequently remembered versus forgotten compound words, i.e., parietal SME, is consistent with both a P300 and a P600 view. However, a P600 interpretation of the parietal SME fits well with schema-based learning as it might be a direct correlate of the integration of the word constituents into the schema representation. To conclude, both presented explanations might account for the parietal SME but the contributions of the processes underlying the P300 and the P600 to the parietal SME cannot be disentangled in the current study and should be addressed in future studies. Moreover, these approaches are not mutually exclusive, as there is an ongoing debate in psycholinguistics whether the family of P600 positivities belongs to the wider P3 family (Friederici et al., 2001; Osterhout et al., 1996). According to a recent account, the P600 might mark the "point in time where a linguistic entity has achieved subjective significance and some form of adaption process is underway" (Sassenhagen et al., 2014, p. 37).

However, these considerations are beyond the scope of this article. Another important topic, which should be addressed in future research, is how the discussed processes reflected in the parietal SME relate to neuroanatomical models of schema-based learning. Even though inferences from scalp ERPs on underlying brain structures are difficult to draw, it is tempting to speculate that the parietal SME, consistently found when schema-congruent information is successfully encoded, is an electrophysiological correlate of the lower mPFC activity and/or the weaker connectivity between the mPFC and the hippocampus, observed in brain imaging studies when schema-congruent information is encoded (van Kesteren et al., 2010). The late frontal subsequent memory effect While an SME in the 700 to 900 ms time interval was present only in the congruent condition, in a still later time interval (900–1200 ms), more positive going waveforms for hits than for misses were obtained irrespective of the encoding condition. Kamp et al. (2017) reported a similarly late and frontally distributed SME that was not affected by encoding condition. With its clear frontal topography, this effect differs clearly from the preceding parietal SME. Moreover, as it was indistinguishable between the two encoding conditions, it presumably reflects processes in brain networks that contribute equally to successful encoding in both conditions. As the late and frontally distributed SME has frequently been observed when relations between arbitrary items had to be encoded (Kamp et al., 2017; Karis et al., 1984; Mecklinger & Müller, 1996), it has been taken as an index of more general successful inter-item encoding. Alternatively, these late effects could reflect post-encoding mnemonic processing like the reactivation of memory traces or the transformation of working memory representations in long-term memory (Cohen et al., 2015).

4.6 Conclusion

In the present study, we investigated the mechanisms by which prior semantic knowledge facilitates episodic encoding of new information and extend the schema framework to the learning of novel compound words. We found superior associative memory performance in the schema-congruent context condition, as compared to the neutral condition. Analyses of event-related potentials revealed an N400 effect that extended in a post-N400 time interval

(500 to 700 ms). The N400-SME and the post-N400 SME did not differ across conditions and this pattern of results is interpreted in that the processes reflected in the N400-SME reflect semantic integration of the pre-activated concepts of the constituents, irrespective of context. In a later time-interval, an SME with a parietal distribution was larger for words preceded by a congruent context. This effect, which we link to the schema-supported formation of a conceptual compound representation, could also account for the superior memory performance in the congruent context condition. An additional late frontal SME was present in both conditions and might reflect inter-item binding of the underlying compound word constituents and their respective fictional contexts (Kamp et al., 2017).

4.7 Next steps

In the current work, we were interested in whether schema-based learning of novel associations is performed by a specific learning mechanism that might include a unitization process. So far, we found first evidence for special mechanisms underlying schema-based encoding of novel associations. In particular, this mechanism includes processes indicated by the parietal subsequent memory effect that is larger in the schema congruent as compared to the neutral condition. We linked this effect to the schema-supported formation of a conceptual compound representation which might be unitized in nature. This idea is supported by research from Kamp and colleagues (2017) in which a similar parietal SME was found being larger for a unitization as compared to a non-unitization control condition. However, data during encoding can only provide indirect evidence for the formation of unitized representations. Thus, in a next step, ERP data during the recognition memory test are considered. In addition, we were interested in the occurrence of a specific behavioral data pattern that should indicate the formation of unitized representations. Hereby, associative memory performance should benefit more from unitization than item memory performance (Parks & Yonelinas, 2015). We tested this idea in the next step.

5 Schema-Congruency Supports Familiarity-Based Retrieval of Novel Compound Words (Exp. 1)[5].

5.1 Introduction

Do you know the compound word *awe walk*? And if yes, does the word only feel familiar and you maybe know its meaning, or do you remember the exact episode when you first heard of it? The quality of your memory probably depends on several factors, for example how special the situation of its first occurrence was (did you have a heated discussion with your colleague, or did you just superficially read it in the media) and how often you already heard and used the word. Another important factor influencing if and how you remember the word might be how well you can make use of your prior world knowledge of the underlying concepts. If you know what *awe* and *walk* mean, you can integrate both constituents into the novel whole word meaning, i.e., walks during which one focuses on awe (see Keltner & Haidt, 2003), one's sense of wonder, and one goes to new places (Sturm et al., 2020). This might be different for a word like *elderberry,* as it is not clear what *elder* contributes to the whole-word meaning. Consequently, you may not be able to embed both constituents into the whole word representation in this case. An interesting and not yet explored question is how a congruent definition that is provided for a new compound (e.g., *walks during which one focuses on awe*), affects the way in which such novel compound words are remembered.

Traditionally, memory research converged on the idea that learning of associations relies on the hippocampus (e.g., Davachi, 2006), in which respective memory representations are initially stored. Only after time, storage is thought to be independent from the hippocampus, relying more on neocortical areas (i.e., systems consolidation, Squire & Alvarez, 1995; Dudai, 2012; see Gilboa & Moscovitch, 2021, for a recent review). However, in reviewing recent research, Hebscher and colleagues (2019) put forward the idea that new cortical engrams may be rapidly viable if they can be associated with active, well-

[5] This chapter consists of a modified version of a submitted manuscript (Meßmer et al., submitted) in which paragraphs have been removed or slightly edited to avoid redundancy.

established cortical engrams. One factor promoting this relationship is relatedness to prior knowledge and the idea of a prior-knowledge-driven, distinct learning mechanism dovetails with recent neuroscientific theories on schema-based learning (e.g., van Kesteren et al., 2012; Gilboa & Marlatte, 2017).

Memory schemas, originally introduced to memory research by Bartlett (1932), denote “higher-level knowledge structures that organize lower-level representations from long-term memory” (Gilboa & Marlatte, 2017, p.618). A schema acts as a template for online information processing by enabling the use of already existing (semantic) knowledge. More importantly, it also affects which aspects of events are encoded and retained in memory or later forgotten (i.e., Pichert & Anderson, 1977; see Bartlett, 1932; Gilboa & Marlatte, 2017). Hereby, memory for to-be-learned schema congruent information is better than memory for neutral or incongruent information (e.g., Höltje et al., 2019; Meßmer et al., 2021; Schulman, 1974; Pichert & Anderson, 1977). However, memory for schema-incongruent information (similar to schema-congruent information) can also be better than memory for neutral information (Greve et al. 2019). Furthermore, prior knowledge has been found to influence episodic encoding of single items (e.g., Höltje et al., 2019) and of associations between pre-experimentally unrelated items (e.g., Meßmer et al., 2021; Staresina et al., 2009). An interesting question is how schema-supported encoding of novel associations affects the retrieval of these associations and how this is reflected in the neural correlates of memory retrieval.

In her process-based perspective on memory systems, Henke (2010) proposes that rapid non-hippocampal learning for associations is possible when associations are unitized. Unitized representations of associations are the result of a unitization process, by which to-be-associated items are integrated and then represented as a single entity (Graf & Schacter, 1989). Combining both approaches leads to the question whether prior knowledge-based learning supports associative memory through the formation of unitized representations. With the current study, we aim to address this question by providing insights into the neurocognitive mechanisms underlying associative recognition of novel compound words, depending on whether or not its learning was supported by a preceding congruent definition.

The neurocognitive mechanisms during memory retrieval depend upon the type of memory, which is probed. One such type is recognition memory, i.e., memory for the previous occurrence of an event (e.g., *awe walk*; Mandler, 1980). In dual-process models of recognition memory (Yonelinas, 2002; Yonelinas et al., 2010), familiarity is defined as a fast-acting strength process, subjectively perceived as a feeling of knowing (e.g., "I know the word *awe walk*"; Mandler, 1980; Yonelinas, 2002), whereby recollection denotes the retrieval of specific details of a study event (e.g., "My colleague showed me this interesting article about *awe walks*"; Yonelinas, 2002).

One way to estimate the contribution of familiarity and recollection to recognition judgements is the use of event-related potentials (ERPs). ERPs allow to monitor memory processes online and to dissociate neurocognitive processes with high temporal resolution. In recognition memory tasks, ERPs are typically contrasted between stimuli correctly identified as "old", and correctly rejected as "new". This old/new effect serves as an index of general memory success (Friedman & Johnson, 2000), and allows to compare memory-related ERP components across different recognition memory studies. Hereby, two such old/new effects are of special interest. A late, left-parietal old/new effect between 500–700ms after stimulus onset has been interpreted as reflecting recollection (Rugg & Curran, 2007). The putative correlate of familiarity, the FN400, onsets earlier (300 - 500ms after word onset) and has a frontally focused distribution (Woodruff et al., 2006; Rugg & Curran, 2007; see Paller et al. 2007, for a different view). Importantly, whilst familiarity is relevant for item recognition (Mandler, 1980), it can also contribute to associative recognition judgements when unitized representations have been formed (e.g., Bader et al., 2010; Henke, 2010; Parks & Yonelinas, 2015).

In an illustrative study, Bader et al. (2010) presented participants with pre-experimentally unrelated word pairs, either together with a definition enabling their processing as a new compound word (unitization condition) or with a sentence in which the two words had to be filled in and were thus processed as separate items (control condition). In a subsequent (surprising) associative recognition memory test, participants had to indicate whether a word pair was intact, recombined, or new. Consistent with the view that unitized representations show less dependence on recollection but rather support familiarity-based recognition, they found ERP evidence for a higher contribution of familiarity-based recognition after unitization encoding, indicated by an early old/new effect, whereas the late parietal old/new effect, the putative correlate of recollection, was dominant in the control condition. Notably, in contrast to the typical mid-frontal distribution of the FN400, the early old/new effect in this study showed a posterior distribution, which was interpreted as an N400 (attenuation) effect, reflecting the larger conceptual fluency for unitized compound words than for unstudied word pairs.

In line with the idea that familiarity is not a single construct (e.g., Bridger et al., 2014), this N400 (attenuation) effect has been interpreted as reflecting absolute (baseline) familiarity (e.g., how familiar you would have rated awe walk before reading this paper; Mecklinger & Bader, 2020). In contrast, in a typical recognition memory experiment with pre-experimentally known items, relative familiarity contributes to recognition judgements (If you already knew awe walks before reading this paper, its additional occurrence makes this word relatively more familiar than what you would expect for this word based on its absolute familiarity; Mandler, 1980). In a recent neurocognitive model of recognition memory, Mecklinger and Bader (2020) propose that relative and absolute familiarity are two independent but interwoven mechanisms. While relative familiarity tracks the increment of a recent to a present exposure of an event, absolute familiarity can be diagnostic for recognition memory judgements for novel stimuli whenever those novel words are assigned a meaning in a learning situation. Acquiring a meaning for a novel compound word results in an absolute familiarity signal for this compound word which is sufficiently diagnostic to distinguish a studied word from an unstudied novel word for which no meaning is assigned (Mecklinger & Bader, 2020).

Thus, in the current study we combined the idea that new cortical engrams may be rapidly viable if the respective information relates to prior knowledge (Hebscher et al., 2019) with the approach of fast non-hippocampal associative learning for unitized associations (Henke, 2010). Hereby, our goal was to investigate whether newly learnt compound words can be unitized when the words comprising the to-be-learned compound are congruent with the definition given for this compound. To address this question, we manipulated schema congruency between a fictional definition and a novel compound word. Each participant is presented with each compound word with either its congruent or its neutral definition, with both, the neutral and a congruent context being presented across participants. In a subsequent associative memory test, participants were presented with intact compound words from the prior learning phase, recombined compound words, for which both constituents had been presented in the learning phase but within different compound words and completely new, i.e., yet unpresented, compound words. Their task was to classify each compound word as being intact, recombined or new. While a majority of studies exploring the circumstances under which unitized representations can be formed manipulate the congruency between two items (e.g., Ahmad & Hockley, 2014; Diana et al., 2011; Kriukova et al., 2013; Rhodes & Donaldson, 2007; Tibon et al., 2014), our approach was to manipulate the congruency between the to-be-associated words and a preceding fictional definition.

If schema-based learning contributes to associative learning in this way, we expected a congruency effect to occur (Bein et al., 2014; 2015), i.e., better associative memory performance in the congruent as compared to the neutral condition. If novel compound words are learnt in this way, this should enhance their absolute familiarity and ensuing recognition memory decisions should be given on the basis of absolute familiarity. In line with this, we expect an N400 attenuation effect (the ERP correlate of absolute familiarity) for words learned in a congruent but not in a neutral context.

Additionally, following Parks and Yonelinas (2015), we assumed that evidence for unitization can also be derived from a comparison of item and associative memory. If unitization enhances familiarity specifically for associations (and not for single items) then there should be larger benefits for associative memory than for item memory in the (unitization supporting) congruent condition. Conversely, if congruency leads to similar benefits for item and associative memory this would imply that a common mechanism such as deep encoding or semantic elaboration may have boosted memory performance in both tests comparably.

5.2 Methods

5.2.1 Participants

N = 43 young adults volunteered for this study, having been recruited via flyers and local databases. The sample size estimation was conducted based on expected ERP effects in the learning phase (see Meßmer et al., 2021; chapter 4). For the current analyses, data from *n* = 10 participants had to be excluded due to failures during recording (*n* = 2), because the stimulus materials were known from another study (*n* = 1), because they reported that they intentionally studied the stimuli or did not give an indication (*n* = 5) or did not provide more than or equal to 10 artifact-free trials (*n* = 2). Thus, the final sample consisted of *N* = 33 participants (22 females, 11 males, with an age range from 18 to 31, *Mdn* = 23 years, *SD* = 3.72). All participants performed significantly above chance level, which was verified with a binomial test ($p < .05$). All participants were students of Saarland University or volunteers from the community and reported being in good health, not suffering from any neurological or psychiatric conditions and having normal or corrected-to-normal vision. Further, all participants were right-handed, as assessed with the Oldfield Handedness Inventory (Oldfield, 1971), and reported being native speakers of German. Participants gave their informed consent and were reimbursed with 10E/h. Participants were debriefed after the experiment. The experiment was approved by the ethics committee of the Deutsche Gesellschaft für Sprachwissenschaft (#2017-07-180423).

5.2.2 Stimulus Materials

For this study, stimulus materials comprised 240 novel compound words, each consisting of two unrelated nouns (e.g., casinoshepherd), together with a congruent and a neutral definition.

Example:

"Ein Angestellter, der eine Spielbank bewacht, heißt… Kasinohirte" (congruent)

(An employee who guards a gambling house is called… Casinoshepherd)

"Ein Angestellter, der eine Markise bewacht, heißt… Kasinohirte"
(neutral)

(An employee who guards an awning is called… Casinoshepherd)

A definition was stated congruent when it reasonably explained how the two nouns could be combined to a new concept. In a rating study with an independent sample of participants, the congruent definition was rated as significantly better explaining how both constituents form the novel concept than the neutral definition (see Meßmer et al., 2021; see 3.2.3). The three dots following the definition in the example are for illustrative purposes, only, and were not shown in the experiment.

For the memory test, we additionally selected 120 recombined compound words, as well as 80 new compound words (see Meßmer et al., 2021; chapter 4). New compound words consisted of two unrelated nouns, which were not used elsewhere in the material. Recombined compound words were included to ensure that participants would not be able to solve the task by using item recognition alone and constructed by newly combining the modifier and the head of two different compound words. Hereby, only one out of two possible combinations were used. It was assured that the nouns still were semantically unrelated.

The selected stimuli were divided into two sets (Set 1 and Set 2), consisting of 120 compound words each. Two encoding lists were created, whereby for the first list, compound words of Set 1 were presented with a congruent context and compound words of Set 2 were presented with a neutral context. This assignment was reversed for the second list. Which encoding list was used varied across participants, whereby both lists were presented approximately equally often.

For the test phase, stimuli were further divided into four subsets of 60 compound words, each, by halving Set 1 and Set 2, respectively. Compounds of each subset were presented as intact for one part of the participants and as recombined for the other part of the participants, so that when compound words of Set 1a and Set 2a were presented as intact compound words, the other half (Set 1b and Set 2b) was presented as recombined compound words, and vice versa. Each test list consisted of 120 intact compound words, 60 recombined compound words, and 80 new (yet unpresented) compound words. The new compound words were identical for each participant. Due to the exclusion of datasets mentioned above, the lists are not fully counterbalanced. However, each of the four combinations of encoding lists and test lists occurred approximately equally often.

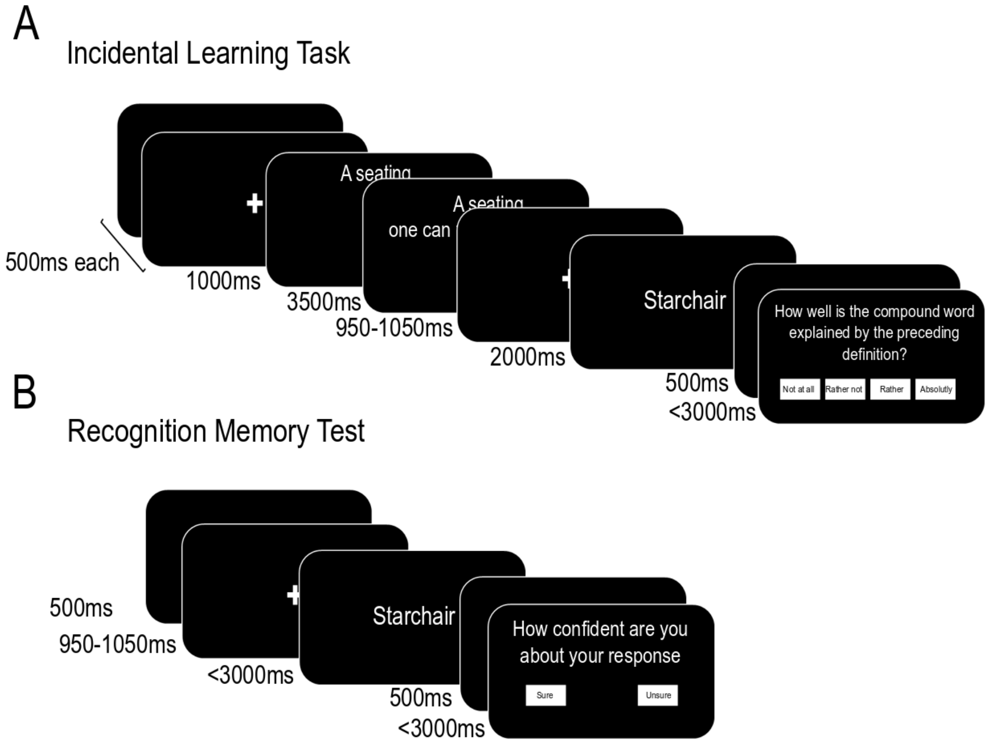
5.2.3 Procedure

After having given their written–informed consent, participants completed several questionnaires, one about their general health, one about demographic aspects and the Oldfield Handedness Inventory (Oldfield, 1971). Next, electroencephalography (EEG) was applied, and participants were sat in a dimly lit, sound-absorbing chamber.

The experiment was created using E Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). The experiment proper consisted of an incidental encoding phase, a retention interval with a duration of 10 minutes and a test phase. During the encoding phase, participants were presented with 240 definitions, half of them congruent and half of them neutral, followed by the respective novel compound word. Participants were instructed to rate on a scale from 1 (*not at all*) to 4 (*absolutely*) how well the novel compound word denotes the concept given by the definition. Participants responded on a keyboard by using the keys x, c, n, and m with their index and middle fingers of each hand. The trial procedure can be seen from Figure 10. For a more detailed description of the trial procedures of the learning phase, we refer the reader to Meßmer et al. (2021; see chapter 4.3.3). Before the learning phase, participants completed eight practice trials to familiarize with the task.

Figure 10

Trial Procedures of Learning and Test Phase



Note. Panel A depicts the trial procedures of the incidental learning task. Panel B shows the trial procedures of the recognition memory test. The English translations of the stimulus materials are for illustrative purposes, only, as all stimuli were presented in German.

The learning phase was followed by a 10-minute retention interval. During this interval, participants performed two distractor tasks. At first, an adapted computerized version of the Digit Symbol Task (Wechsler, 1955) from Häuser et al. (2019) was performed for approximately 5 minutes, followed by 2.5 minutes of backwards counting in steps of 3. Only then, participants were told about the upcoming test phase. During the test phase, participants were presented with one of the two test list versions, consisting of 120 intact compound words, 60 recombined compound words and 80 new, i.e., yet unpresented compound words. A trial started with a continuously jittered fixation cross (950-1050 ms). Then, the compound word was presented (for up to 3000 ms), until participants gave their response. Participants gave their answer on a keyboard by using the keys f, j, and k to indicate if the compound word was intact, recombined or new. Key assignment was varied by using a latin-square design, ensuring that across participants, each response option was used with similar frequency. After a 500 ms blank screen, participants were asked to indicate their confidence on the previous response (*sure* or *unsure*) using their index fingers, whereby key assignment was ascending for a part of the participants and descending for the other part of participants. The confidence scale remained on the screen for up to 3000 ms or until participants gave their response and was only presented if a response had been logged on the compound word. A trial ended with a blank screen, which was presented for 500 ms.

Stimulus presentation was pseudo-randomized for the encoding and test phase, with the limitation of not more than 3 consecutive trials in the same context condition (encoding phase) or not more than 3 consecutive trials requiring the same response (test phase). In both phases, there were self-paced breaks after 60 trials (encoding phase) or 65 trials (test phase), respectively.

5.2.4 Data Acquisition and Pre-Processing

The EEG was continuously recorded from 28 Ag/AgCl scalp electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC3, FCz, FC4, FC6, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P7, P3, Pz, P4, P8, O1, O2, and A2), using BrainVision Recorder 1.0 (Brain Products, Gilching, Germany), whereby all electrodes except from A2 were embedded in an elastic cap (Easycap, Hersching, Germany). Electrode positions followed the extended international 10-20 system (Jasper, 1958). AFz was chosen as ground electrode and two additional electrodes were applied on the left (A1) and right (A2) mastoid, respectively. The signal was online referenced to the left mastoid electrode (A1) with the exception of one participant for whom some eye and mastoid electrode channels had been interchanged by mistake. For this dataset, data were online referenced to the left canthus electrode. Channel assignment was corrected offline and in an additional step, data were re-referenced to left mastoid so that before the actual pre-processing, all datasets had the same reference. Electroocular activity was assessed via four additional electrodes, which were placed above and below the right eye and outside the outer canthi of both eyes. All electrode impedances were kept below 5 kOhm with the exception of the electroocular electrodes' impedances. Data were sampled at 500 Hz. An online filter from 0.016 Hz (time constant 10 s) to 250 Hz was applied.

Offline, the data were pre-processed using the EEGLAB (version 2019.1; Delorme & Makeig, 2004) and ERPLAB (version 7.0; Lopez- Calderon & Luck, 2014) toolboxes for MATLAB (MathWorks, Inc.). First, the data were down sampled to 250 Hz and re-referenced to the average of the left and right mastoid. Thereafter, data were filtered with a second-order Butterworth bandpass-filter from 0.05 Hz to 30 Hz (-6dB half-amplitude cutoff, with DC removal) and 50 Hz powerline fluctuations were removed with a Parks- McClellan notch filter (default setting order 180; with DC removal). Data were pre-segmented by discarding all data points exceeding a time period from 1000 ms before a stimulus onset marker to 2500 ms after a stimulus onset marker. Then, bad segments and experimental breaks, as well as practice trials, were manually discarded. Thereafter, the weights and sphere matrix of an independent component analysis (ICA)[6] run with the infomax algorithm *runica* ICA were applied, and components associated with eye movements and muscular artifacts were identified and removed (up to 5 components per participant). Data were then segmented into epochs of 1696 ms around compound word onset, including a 200 ms baseline. Following baseline correction, a semi-automatic artifact rejection was applied, using the following criteria: a maximally allowed amplitude of -75 up to 75 µV, a maximal difference of values of 100 µV during intervals of 200 ms (window steps of 100 ms), a maximally allowed voltage step of 50 µV/s and a maximum of 200 ms of sample points with a deviation from -.5 to .5 µV from the maximum voltage in this epoch.

5.2.5 Data Analysis

Similar to Kamp et al. (2016), we included the definition fit rating of the learning phase as an additional constraint to both, behavioral data analysis and ERPs of the test phase, as although definitions were created (and rated) to map on the congruent vs. neutral definition distinction, there might be participant-specific variations of definition goodness within conditions (Wiegand et al., 2010). Therefore, hits to intact compound words in the congruent condition were

[6] The ICA was run with a more conservative second-order Butterworth bandpass-filter from 0.5 Hz to 30 Hz (-6dB half-amplitude cutoff, with DC removal) to optimize performance.

only included if they were rated with a 3 *(rather)* or 4 *(absolutely)* on the 4-point scale in the initial learning phase, whereby inclusion of hits for intact compound words in the neutral condition required a rating of 1 *(not at all)* or 2 *(rather not)*. To calculate average hits, *M* = 32.06 trials (*SD* 9.36, *range* 11–49) were used in the congruent condition and *M* = 23.46 trials (*SD* 7.53, *range* 11–41) were used in the neutral condition. Correct rejections were based on *M* = 49.12 trials (*SD* 13.01, *range* 15–71).

For all analyses, the significance criterion of $p < .05$ was applied. Data were analyzed using R (version 3.6.1; R Core Team, 2019) and RStudio (Version 1.2.5001; RStudio Team, 2019) and IBM SPSS statistics (version 26). Whenever non-hypothesis-driven multiple testing was required, the Bonferroni-Holm correction (Holm, 1979) was applied. The reported corrected *p*-values were calculated with the function *p.adjust* of the R package *stats* (R Core Team, 2019).

Associative memory performance was estimated by the measure Pr_{As}, (hits-false alarms), the difference between the probability of an intact response to an intact compound word (hits) and the probability of an intact response to a recombined compound word (false alarms; see de Chastelaine et al., 2016, and Huffer et al., 2022, for similar procedures). In this metric, the false alarm rate includes recombined pairs mistakenly recognized as intact relative to all recombined pairs with at least correct item memory (recombined pairs as recombined and recombined pairs as intact). This false alarm rate was subtracted from the hit rate, which was calculated as the proportion of intact pairs correctly recognized as intact relative to all intact pairs with at least correct item memory (intact pairs as intact and intact pairs as recombined). To examine whether the congruency manipulation affects associative and item memory performance in the same way, an additional metric for item memory (Pr_{It}) was calculated. Item memory refers to the ability to generally discriminate pairs containing old constituents (items) (i.e., intact and recombined pairs) from pairs containing new items, i.e., new pairs. Accordingly, the item hit rate was calculated as the sum of all intact or recombined items classified as either intact or recombined, divided by the sum of all intact and recombined trials. The item false alarm rate was calculated as the sum of all new compound words classified as intact or recombined, divided by the number of new items, i.e., 80. Behavioral outliers were defined as extreme values, i.e., with a standardized *z*-value greater than 3.29 above or smaller than -3.29 below the mean (Field, 2009, p. 179). One dataset was detected as a behavioral outlier. We calculated all behavioral analyses once with and once without this dataset, with no qualitative differences between the respective results. Therefore, we report the results with this dataset included.

As sphericity is usually violated in EEG data, we used the multivariate approach of repeated measure analysis of variance (MANOVA) for ERP analyses, which is more robust against such violations of sphericity (Dien & Santuzzi, 2005; Picton et al., 2000). Significant effects are further explored in follow-up MANOVAs and paired-samples *t* tests. As measures of effect size, we report Hedges's g_{av} for effects from paired-samples *t* tests with the formula provided in the spreadsheet (Version 5) from Lakens (2013) and Pillai's trace, which is identical to partial eta-squared (η^2), for multivariate analyses of variance (MANOVAs), respectively.

Following a recommendation of Luck & Gaspelin (2017), to avoid unnecessary accumulation of type I errors, we eliminated all factors that were not relevant for our hypotheses from our MANOVA designs. Hence, all analyses were restricted to a subset of electrodes relevant for the respective analyses. As we expected the N400 attenuation effect as a measure of absolute familiarity to be broadly distributed across the scalp (Bader et al., 2010; Wiegand et al., 2010; see Mecklinger & Bader, 2020 for a review), we chose a broad array of electrodes (i.e., F3, Fz, F4, C3, Cz, C4, P3, Pz, P4) for the analyses of this effect. To quantify the late parietal old/new effect, typically showing a left lateralization for language materials (Rugg & Curran, 2007), we chose the electrodes P3, Pz, P4 to statistically confirm the left lateralization. In addition, all topographical analyses were conducted with the electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, P4.

Due to the temporal extension of the N400, we aimed to define two N400 (absolute familiarity) time windows of 200ms duration, each, post hoc. Those should be centered around the two N400 peaks for correct rejections at electrodes Cz, in the two time windows from 300-500 ms and 500-700 ms. Those are at 464 ms and 576 ms, respectively. However, this approach would have resulted in two overlapping N400 time windows. Thus, to cover the whole variance in the vicinity of the two N400 peaks in two non-overlapping time windows, we took the arithmetic mean between both peaks (i.e., 520 ms) and defined the first time window from 364 to 520 ms and the second time window from 520 to 676ms. To obtain an independent and non-overlapping time window for the late parietal old/new effect, indicative of recollection (e.g., Rugg & Curran, 2007), a third adjacent time window from 676-876 ms was chosen.

A two-step procedure was used for the analysis of the ERP effects. First, we tested whether there are statistically significant old/new effects in both conditions. For this purpose, we ran a MANOVA with the factors type (intact hit, correct rejection), anteriority (frontal, central, parietal) and laterality (left, middle, right) for each condition in both N400 (absolute familiarity) time windows. For the late parietal old/new effect, the correlate of recollection, a MANOVA with the factors type (intact hit, correct rejection) and laterality (left, middle, right) was calculated for parietal electrodes for each condition. In a next step, we were interested in whether old/new effects differ across congruency conditions. As the ERPs to correct rejections are condition-unspecific, the critical comparison to test whether old/new effects differ across congruency conditions is to contrast the ERPs to congruent and neutral hits. Thus, we tested whether there are condition-specific differences between intact hit ERPs using a MANOVA with the factors congruency (congruent, neutral), anteriority (frontal, central, parietal) and laterality (left, middle, right) for the first and the second N400 (absolute familiarity) time window. For the third time window, we conducted a MANOVA with the factors congruency (congruent, neutral) and laterality (left, middle, right) at parietal electrodes at which the late parietal component (LPC) is largest. For the sake of readability, we only report significant effects including the factors congruency or type.

5.3 Results

5.3.1 Behavioral Results

The behavioral data for both, item and associative memory performance are illustrated in Table 2.

Table 2

Behavioral data

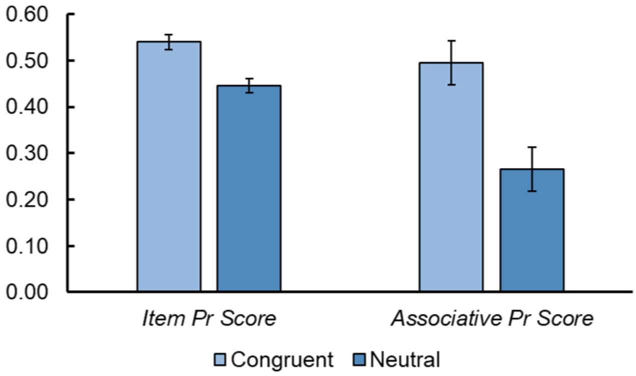
	Associative memory		Item memory	
	congruent	neutral	congruent	neutral
Pr score	.50 (.24)	.27 (.21)	.54 (.16)	.45 (.15)
Hit rate	.81 (.12)	.61 (.14)	.86 (.08)	.76 (.11)
FA rate	.32 (.21)	.34 (.21)	.32 (.15)	

Note. Standard deviations are given in parentheses.

In line with our hypothesis, associative memory performance, indicated by associative *Pr*, was higher in the congruent than in the neutral condition, $t(32) = 4.88$, $p < .001$, $g_{av} = 1.00$ (one-sided $M_{congruent} = .50$, $SD = .24$, $M_{neutral} = .27$, $SD = .21$; see Figure 11). To explore the contribution of hits and false alarms to this condition difference, we calculated a Congruency (congruent, neutral) x Type (hit, false alarm) - MANOVA. This analysis revealed a main effect of congruency, $Pillai = .31$, $F(1, 32) = 14.05$, $p = .001$, a main effect of type, $Pillai = .83$, $F(1, 32) = 151.22$, $p < .001$, and a significant interaction, $Pillai = .43$, $F(1, 32) = 23.81$, $p < .001$. Further examination revealed that only hit rates differed significantly across conditions, $t(32) = 10.64$ $p < .001$, $g_{av} = 1.51$ (one-sided; $M_{congruent} = .81$, $SD = .12$, $M_{neutral} = .61$, $SD = .14$), whereas no such difference was found for false alarm rates, $t(32) = -0.63$, $p = .534$, $g_{av} = 0.13$ ($M_{congruent} = .32$, $SD = .21$, $M_{neutral} = .34$, $SD = .21$).

Figure 11

Mean Item and Associative Pr Scores.



Note. The error bar depicts the standard error of the mean difference, obtained for the congruent versus neutral comparison for the item and the associative *Pr* score, respectively.

To address the idea that associative memory benefits more than item memory from the schema congruency-driven formation of unitized representations, we calculated a Congruency (congruent, neutral) x Memory Type (Item, Association) – MANOVA on *Pr* Scores. This analysis revealed a main effect of congruency, *Pillai* = .57, $F(1, 32) = 42.39$, $p < .001$, a main effect of memory type, *Pillai* = .29, $F(1, 32) = 12.80$, $p = .001$, and a significant interaction, *Pillai* = .19, $F(1, 32) = 7.50$, $p = .010$, whereby the congruency effect was larger for associative *Pr* scores (see above) than for item *Pr* scores, albeit there was also a congruency effect on item *Pr* scores, $t(32) = 6.09$, $p < .001$, $g_{av} = 0.58$ ($M_{congruent} = .54$, $SD = .16$, $M_{neutral} = .45$, $SD = .15$; see Figure 2).

5.3.2 ERP Results

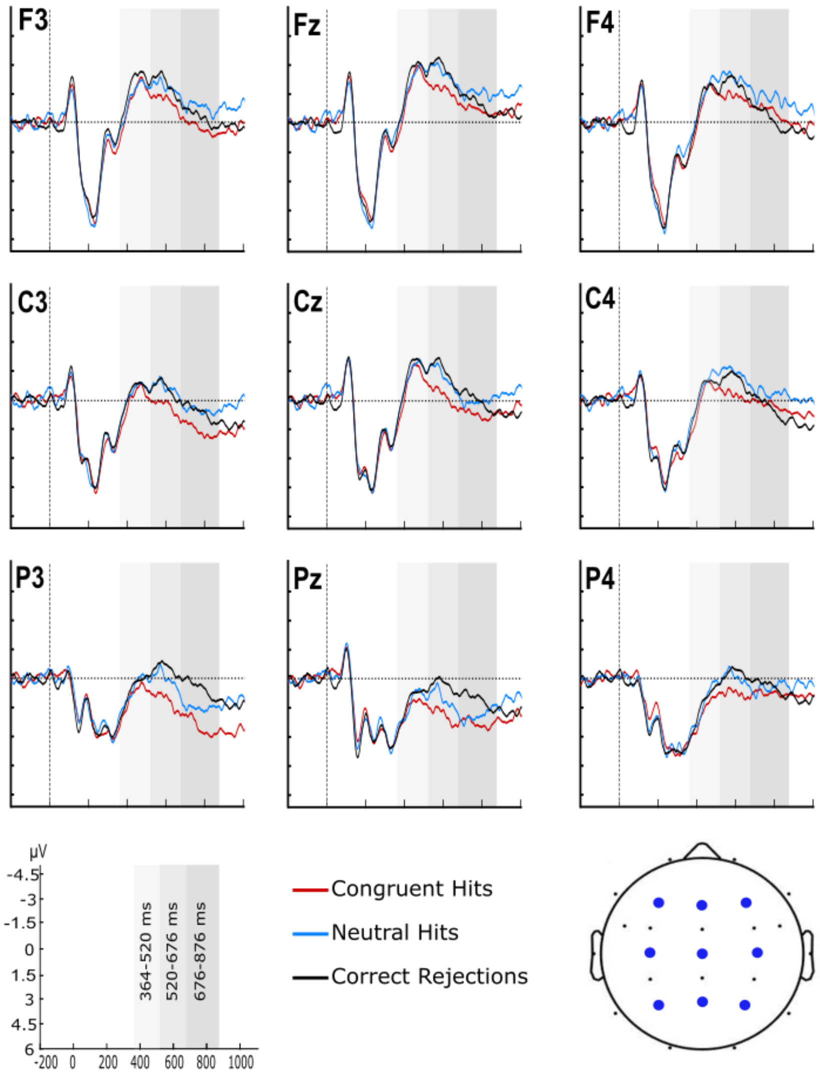
Figure 12 depicts the grand average ERP waveforms elicited by the compound words in the test phase. In order to investigate differences in absolute familiarity between conditions, we had planned to analyze variations in the N400 within a time window from 300 to 500 ms (Mecklinger & Bader, 2020). However, as especially evident in the ERPs to correctly rejected new compound words, the N400 appears to comprise the already mentioned two consecutive peaks and extends approximately 200ms beyond the 300 to 500 ms time window. Whilst waveforms for correctly rejected new compounds and hits in the neutral condition are highly similar between 350 and 700 ms, the N400 for congruent hits is attenuated in the second half of this prolonged N400 time window. This effect is broadly distributed across the scalp and present at all electrodes. This N400 attenuation effect is followed by a later old/new effect with a more left-lateralized scalp topography in both conditions, which is present between 700 and 900 ms. The late old/new effect takes the form of more positive ERPs for hits than correct rejections. These observations were examined in a series of statistical analyses.

5.3.2.1 Old/new effects in the congruent condition

A Type x Anteriority x Laterality-MANOVA revealed no significant effects including the type factor in the first time window (all $p_s > .247$). The Type x Anteriority x Laterality-MANOVA in the second time window revealed a significant main effect of type, *Pillai* = .26, $F(1, 32) = 11.45$, $p = .002$ and a significant Type x Laterality-Interaction, *Pillai* = .25, $F(2, 31) = 5.07$, $p = .012$. To resolve the Type x Laterality-Interaction, we averaged mean amplitudes across the anteriority factor and calculated Bonferroni-Holm-corrected paired-samples *t* tests, comparing hits and correct rejections for left, middle and right electrode sites, separately. These revealed significant differences with more positive waveforms for hits versus correct rejections at left, middle and right clusters, whereby effect sizes are larger at left and middle sites than at right sites (left: $t(32) = 3.66$, $p = .003$, $g_{av} = 0.26$; middle: $t(32) = 3.58$, $p = .003$, $g_{av} = 0.27$; right: $t(32) = 2.48$, $p = .019$, $g_{av} = 0.19$).

Figure 12

ERP waveforms of hits and correct rejections for the nine analyzes electrodes



Note. Hits are intact compound words correctly classified as intact. Correct rejections are new compound words correctly classified as new. Shaded areas indicate analyzed time windows.

The Type x Laterality-MANOVA in the third time window from 676-876 ms revealed a significant main effect of type, *Pillai* = .24, $F(1, 32) = 10.29$, $p = .003$, and a significant interaction between type and laterality, *Pillai* = .46, $F(2, 31) = 13.36$, $p < .001$. Follow-up paired-samples *t* tests for each electrode, separately, revealed a significant old/new-difference at electrode P3, $t(32) = 5.29$, $p < .001$, one-sided, $g_{av} = 0.54$, and Pz, $t(32) = 2.94$, $p = .003$, one-sided, $g_{av} = 0.31$. To sum up, for the congruent condition, we found a broadly distributed absolute familiarity effect in the second time window (520-676 ms) and a left-mid-lateralized recollection effect at parietal recording sites in the third time window (676-876 ms).

5.3.2.2 Old/new effects in the neutral condition

A Type x Anteriority x Laterality-MANOVA in the first time window revealed no significant effects (all p_s > .093). The Type x Anteriority x Laterality-MANOVA in the second time window revealed a significant Type x Anteriority-Interaction, *Pillai* = .24, $F(2, 31) = 4.85$, $p = .015$ and a significant Type x Laterality-Interaction, *Pillai* = .29, $F(2, 31) = 6.30$, $p = .005$. To resolve the significant interaction of laterality and type, mean amplitudes were averaged across the anteriority factor and Bonferroni-Holm-corrected paired-samples *t* tests were calculated for left, middle and right electrode clusters, separately, comparing hits in the neutral condition and correct rejections. These revealed no significant differences (all $p_s > .978$, $g_{av} < .08$). To resolve the significant interaction of anteriority and type, mean amplitudes were averaged across laterality and Bonferroni-Holm-corrected paired-samples *t* tests were calculated for frontal, central and parietal electrode clusters, separately, comparing hits in the neutral condition and correct rejections. Again, no significant differences were found (all $p_s > .618$, $g_{av} < .12$).

The Type x Laterality-MANOVA in the late time window from 676 – 876 ms revealed a significant interaction of type and laterality, *Pillai* = .32, $F(2, 31) = 7.26$, $p = .003$. Follow-up paired-samples *t* tests comparing ERPs to hits and correct rejections at each electrode, separately, revealed a significant effect at electrode P3, $t(32) = 2.26$, $p = .016$, one-sided, $g_{av} = 0.26$. In conclusion, no absolute familiarity effects were found in the neutral condition. There only was a significant left-parietal old/new effect in the third time window.

5.3.2.3 Congruency effects on ERPs to hits

In a next step, we were interested in whether old/new effects differ across congruency conditions. As the intact and recombined word pairs of both conditions were presented together with the same new word pairs, the ERPs to correct rejections were the same in the aforementioned analyses of old/new effects. For this reason, we contrasted the ERPs for hits in both conditions using a MANOVA with the factors congruency, anteriority and laterality. In the first time window, this MANOVA did not reveal any significant effects (all p_s > .437). For the second time window, the Congruency x Anteriority x Laterality - MANOVA yielded a significant main effect of condition, *Pillai* = .14, $F(1, 32) = 5.02$, $p = .032$, reflecting more positive ERPs for hits in the congruent than in the neutral condition, across all electrode sites (see Figure 13A). No other effect including the condition factor was significant.

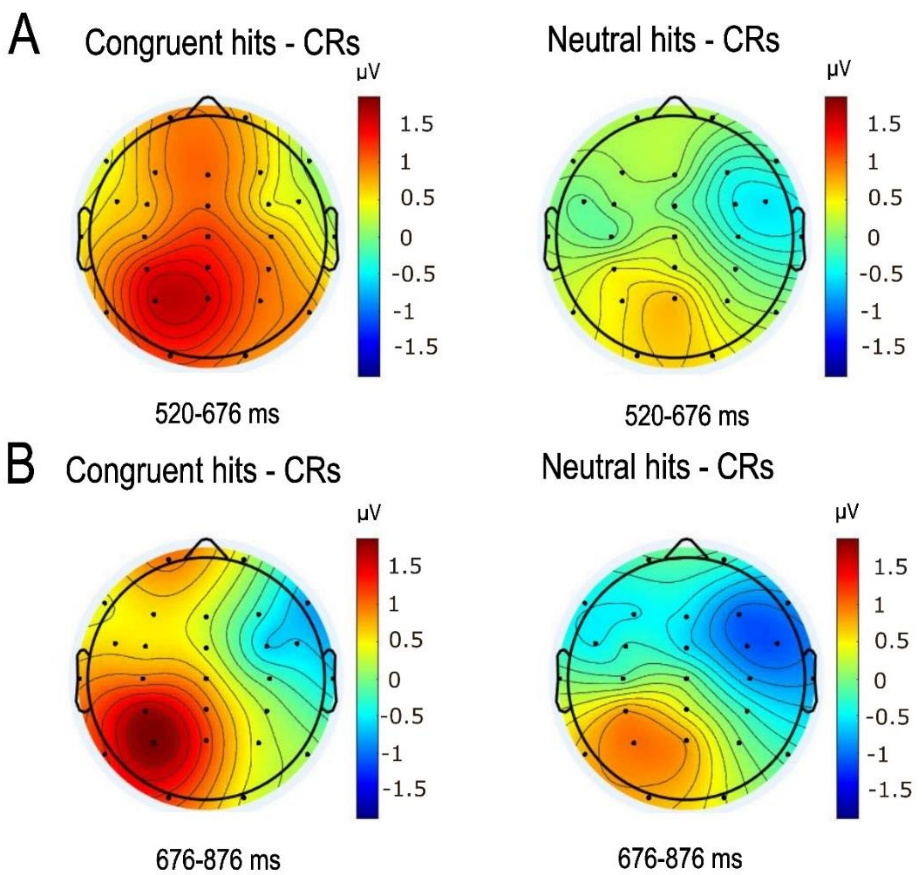
The Congruency x Laterality - MANOVA on parietal electrodes in the third time window from 676-876 ms revealed a significant interaction of congruency and laterality, *Pillai* = .20, $F(2, 31) = 3.95$, $p = .030$. Follow-up, Bonferroni-Holm corrected paired-samples *t* tests calculated at each electrode, separately, revealed a marginally significant effect of congruency at electrode P3, $t(32) = 2.49$, $p = .054$, $g_{av} = 0.25$. Thus, statistically reliable differences between hits in the congruent and the neutral condition were found in the second, absolute familiarity time window, indicating that absolute familiarity differs across congruency conditions. Additionally, there was a marginally significant effect of congruency in the third time window, reflecting a tendency towards a larger contribution of recollection to recognition memory in the congruent as compared to the neutral condition.

5.3.2.4 Topographic profile analysis

The topographic maps contrasting ERPs to hits and correct rejections for each condition are depicted in Figure 13. We compared the topographic profiles of the N400 attenuation effect (520-676 ms) and the late parietal old/new effect (676-876 ms) in the congruent condition to test if both effects differ qualitatively in their topographical distribution, as it would be expected if they reflect functionally different processes (McCarthy & Wood, 1985; Wilding, 2006). Consistent with the view that scaling should not be used for inferences about qualitatively different brain regions (Urbach & Kutas, 2002), in the present context only functional inferences will be made from the scalp topography data (see Wilding, 2006). Thus, we first calculated old/new difference scores for each time window and participant at each recording site. Thereafter, we normalized these data using the vector-scaling method suggested by McCarthy and Wood (1985) and calculated an Effect × Anteriority (frontal, central, parietal) × Laterality (left, middle, right) - MANOVA. This analysis revealed a significant three-way interaction, *Pillai* = .28, $F(4, 29) = 2.77$, $p = .046$, and a significant interaction of laterality and effect, *Pillai* = .53, $F(2, 31) = 17.72$, $p < .001$, supporting the idea that the early old/new effect and the late old/new effect in the congruent condition exhibit distinct topographical distributions and therefore reflect functionally different processes.

Figure 13

Topographic maps



Note. Panel A shows topographic maps (Hits minus Correct Rejections, CRs) in the second time window from 520 to 676 ms for both, the congruent and the neutral condition. Panel B shows the topographic maps (Hits minus Correct Rejections, CRs) in the third time window from 676 to 876 ms for both, the congruent and the neutral condition.

5.4 Discussion

In the present study, our goal was to investigate if schema congruency promotes the formation of unitized representations out of pre-experimentally unrelated associations, and whether these representations can be recognized based on changes in their absolute familiarity. This idea was motivated by new insights on neocortical learning (Hebscher et al., 2019). Here, we combined the idea that new cortical engrams may be rapidly viable if the respective information relates to prior schema knowledge (Hebscher et al., 2019) with the approach of rapid non-hippocampal associative learning for unitized associations (Henke, 2010). From a theoretical point of view, schema-based encoding of novel compound words by means of a congruent definition bears the potential to enable the formation of a conceptual knowledge structure, constituting the whole-word meaning, as well as the information what both words contribute. Hereby, the word representations of the constituents can be linked on a conceptual level, and the newly learnt compound words can be unitized.

5.4.1 Behavioral results

Based on literature on schema-based learning, we expected a typical congruency effect to occur (e.g., Bein et al., 2014; 2015) with better memory performance for schema-congruent than for neutral information. Indeed, associative memory performance, indicated by the associative *Pr* score, was better in the congruent than in the neutral condition. This memory benefit can be traced back to a higher associative hit rate in the congruent than in the neutral condition, as we did not find a condition effect on false alarm rates. Thus, associative memory performance benefited from schema congruency during encoding.

In addition, we explored whether the boost in memory performance in the congruency conditions is larger for associative than for item memory. We presumed that if unitization enhances familiarity specifically for associations (and not for single items), there should be larger benefits for associative memory than for item memory in the (unitization supporting) congruent condition (see Parks & Yonelinas, 2015). This is exactly what we found: the congruency effects were larger for associative than for item memory. Hence, our results support the view that congruency exerts its influence mainly by promoting unitization and not by generally boosting semantic elaboration or deeper encoding (Craik & Tulving, 1975). In the latter case, we would have expected similar benefits from the congruency manipulation for item and associative memory.

5.4.2 ERP results

Consistent with a large number of ERP recognition memory studies (e.g., Bader et al., 2010; Greve et al., 2007; Kriukova et al., 2013; Rhodes & Donaldson, 2007), the main comparison of our ERP analysis was between intact (old) and new word pairs. The ERP results of the current experiment show an N400 attenuation effect (520-676 ms) with a broad, posterior topographical distribution, which is larger for the congruent as compared to the neutral condition. Similar N400 attenuation effects have been taken as a measure of diagnostic fluency in situations in which fluency is salient, as e.g., for novel compound words having a meaning (Bader et al., 2010; Wiegand et al., 2010; see Mecklinger & Bader, 2020). This broadly distributed N400 attenuation effect differs topographically and functionally from the FN400 effect, which is usually found in recognition memory experiments (see Bader & Mecklinger 2017, for a dissociation of both effects in a single experiment). Thus, schema-supported learning of novel compounds increases absolute familiarity for those words. This conclusion is in line with the model proposed by Henke (2010), in which fast learning of novel associations is achieved when these associations are unitized and, familiarity can contribute to recognition memory judgements. Hereby, the absolute familiarity signal is evoked by the assignment of a meaning to the compound word representation, provided by the congruent definition (see Mecklinger & Bader 2020).

Evidence in that schema-congruency promotes familiarity-based remembering of already known items was obtained in a behavioral study by Souza and colleagues (2022). In their second experiment, participants were presented with images of common objects, belonging to one of eight categories. Objects were either shown within a perceptual encoding condition, in which participants had to decide how complex the object is, or in a conceptual schema encoding condition, in which they had to decide if the object belongs to its respective category. Items were either typical or atypical for their category. In a later recognition test, a remember/know/guess procedure was applied (Gardiner, 1988). The authors found that guess responses were more pronounced for typical items, what the authors interpret in that more familiarity contributes to recognition of typical items. Whilst those data provide first evidence for schema-based learning increasing the contribution of familiarity to recognition decisions for known items, in the current study, we were interested in schema-based learning of novel associations, including novel concepts.

Neural evidence for schema-based learning supporting the formation and recognition of unitized representations of novel associations and familiarity-based remembering comes from a brain imaging study by Bader et al. (2014). In this study, unitization was established similar to the present study, by means of a congruent definition. Interestingly, the medial prefrontal cortex, a region known to be critical for schema-congruent encoding (e.g. van Kesteren et al., 2010; see van Kesteren et al., 2012), was found to be more active during recognition of unitized associations than during item recognition and activations in familiarity-related brain regions (in the inferior frontal gyrus and in the parahippocampal gyrus extending into the fusiform gyrus) were selectively found in the congruent definition group. A recent brain imaging study showed that new information (here noun-adjective word pairs) that is initially inconsistent with prior knowledge undergoes a shift from novelty to familiarity in the time course of repeated presentations. In fact, this shift is indicated by convergence, meaning that initially different neural representations of the newly learned word pairs in lateral temporal and frontal brain regions, the medial prefrontal cortex and the medial temporal lobe that changed throughout a few repetitions and gradually became indistinguishable from those of familiar word pairs (Yacoby et al., 2021). This could be interpreted in that rapid neocortical learning, probably underlying both prior knowledge-based and repetition-based learning (see Hebscher et al., 2019), supports familiarity.

An objection against the view that our study supports the notion that newly learned compound words can be unitized and remembered on the basis of familiarity could be that we infer on familiarity from ERP measures alone and did not include additional behavioral measures. However, ERP studies of recognition memory have provided ample evidence that ERP components with specific temporal and topographic characteristics are closely associated with memory processes (such as familiarity and recollection). Some of these studies have validated ERP measures of familiarity by showing similar effects of experimental manipulations on behavioral and corresponding ERP measures (Bruett & Leynes, 2015; Bader et al 2020). Hereby, the FN400 is a reliable and objective measure of familiarity memory, as its amplitude has consistently been found to be sensitive to common operational definitions of relative familiarity (see Rugg & Curran, 2007, for a review). More importantly, N400 (attenuation) effects, indicative for absolute familiarity or conceptual fluency, have been reported in a remarkable number of ERP studies from different laboratories (Strózak et al., 2016; Woollams et al., 2008; Yang et al., 2019; Bruett & Leynes, 2015), adding to the view that ERP components are valid measures of memory processes.

Compared to other studies, the N400 attenuation effect as a correlate of absolute familiarity onsets about 200 ms later in our study and had a slightly longer temporal extension (Bader et al., 2010; Kamp et al., 2016). However, for the following two reasons we think that the interpretation of this effect as a correlate of absolute familiarity is nonetheless plausible. First, the alternative explanation, namely that the N400 attenuation effect reflects early onsetting recollection is highly unlikely, given that topographic profiles of this N400 attenuation effect and the left parietal old/new effect differ qualitatively. Second, a prolonged N400 was obtained for correctly rejected new compound words. This is of relevance, as the N400 to new compound words can serve as a temporal reference for novel compound word processing, independent from the context condition. Because the N400 is prolonged by approximately 180 ms for these new compound words, for which no experimental manipulation was implemented, we are confident that it was the processing of the compound words per se which was prolonged and that this also holds for learned compound words.

The prolonged N400 to compound words might be a consequence of the simultaneous presentation of both compound word constituents, provoking processing of the underlying concepts of both. This is different to most N400 studies, in which the N400 is measured stimulus-locked to a single word (see Kutas & Federmeier, 2011, for a review). Thus, the cumbersome semantic processing of two nouns and their semantic integration might have delayed the processes reflected by the N400.

A question that follows from what we elaborated on before is why the N400 on compound words is not only prolonged relative to N400 elicited by monomorphic words, but also relative to other studies in which novel compound words were learned (e.g., Bader et al., 2010; Wiegand et al., 2010). Critically, in contrast to these studies, compound words in the present study were presented without a space between both constituent words and sometimes contained interfixes. It is possible that the space-free presentation delayed processing of the compounds. Empirical support for this view comes from a study by Inhoff et al. (2000), investigating the effect of spaces versus presentation without spaces between compound constituents. In this study, the concatenative presentation of a compound word without a space in between constituents hampered reading, resulting in increased naming latencies and gaze durations during reading. In light of the temporal delay of the N400 and the missing behavioral evidence for familiarity, our conclusion for the contribution of familiarity to memory for newly learned compound word have to be taken as tentative and need to await a conceptual replication using alternative (behavioral) measures of memory subprocesses.

Interestingly, the ERP absolute familiarity effect in the congruent condition fits well with an idea we proposed earlier in that three interleaved processes[7] underlie the schema-congruency based learning of novel compounds (Meßmer et al., 2021; chapter 4). First, semantic priming of the modifier component by the congruent context results from the additional semantic relationship between the modifier and the context and enables schema-supported processing of the compound word as a whole. Second, context-independent semantic integration of both constituents occurs that depends on lexical characteristics of the constituents. Hereby, a context-independent compound representation can be formed which could then – in a third process - be

[7] Note that presentation order in the text does not correspond to a sequential model. Rather, our data indicate at least partly parallel processing (see Meßmer et al., 2021).

reconciled with the context meaning. Hereby, a conceptually integrated compound representation is formed, linking the new whole-word meaning with prior knowledge as provided by the congruent definition. This integration into prior knowledge structures in turn enables the formation of a unitized, conceptual representation and familiarity-based remembering in an ensuing memory test (Mecklinger & Bader, 2020).

A final issue to be addressed is the relationship between schema congruency and recollection. In contrast to Bader et al. (2010), the late parietal old/new effect, the putative correlate of recollection, shows a tendency towards larger amplitudes in the congruent than in the neutral condition. As this effect is only marginally significant, it should, if at all, be interpreted with caution. However, it might indicate a tendency towards a larger contribution of recollection to recognition judgments in the congruent as compared to the neutral condition. This might be explained by the higher semantic congruency of the contexts in the present study as compared to Bader et al. (2010), in which the relationship between the two words was only established in a non-formalized way. In the current study, schema congruency was achieved by a systematic pattern of semantic relationships. This may have rendered the contexts more memorable enabling recollection of this context information which could then be reflected in the late parietal old/new effect (see also Diana et al., 2007 and Eichenbaum et al., 2007).

5.4.3 Conclusion

In the present study, we investigated the influence of schema-based learning on conceptual unitization of novel compound word associations by using a manipulation of schema congruency.

Similar to previous studies (Bader et al., 2010; Kamp et al., 2016; Wiegand et al., 2010), we found a larger N400 attenuation effect, the ERP correlate of absolute familiarity (Bridger et al., 2014), in the schema-congruent condition, suggesting that newly learned compounds can be unitized when the words comprising the compound are congruent with a prior schema context. In line with a recent neurocognitive model (Mecklinger & Bader, 2020), we interpret this N400 attenuation effect to reflect the contribution of absolute familiarity to recognition judgements.

By providing first evidence that schema-congruency boosts memory by fostering unitization and by enhancing the contribution of absolute familiarity to associative recognition judgements, our results add to brain imaging studies on schema-congruency based learning and remembering (Sommer et al. 2022; Bein et al. 2014; van Kesteren et al., 2012) and shed light on the temporal characteristics underlying the recognition of information which was encoded under the influence of a schema.

5.4.4 Next steps

So far, we found evidence in that schema-congruency benefits the learning of novel compound words with higher associative memory performance for compound words learned with a congruent definition as compared to compound words learned in the neutral definition. Further, ERPs revealed some insights in the neurocognitive processes underlying the schema-congruency effect on associative learning of meaningful concepts, probably resulting in the schema-congruency driven formation of a unitized representation which is recruited in an absolute familiarity process. Henke (2010) and others (Reder et al., 2009) put forward the idea that consciousness does not play a role for how memories are stored. From this, it can be predicted that both explicit and implicit memory tasks can recruit the unitization process in the processing mode model by Henke (2010). This is in line with the neurocognitive account of familiarity we took as a basis in the current work. If conceptual fluency results from the formation of unitized representations, driven by schema congruency and this is what underlies the absolute familiarity effect in experiment 1, we should also be able to observe the consequences of conceptual fluency in an implicit memory task. This prediction was tested in experiment 2 in which we sought to replicate the results of experiment 1 in an implicit memory task. We assumed that conceptual fluency of novel compound words, induced by schema congruency during learning, might be misused in a lexical decision task in which conceptual fluency is diagnostic for deciding if a word is real or not. This should then go hand in hand with higher error rates, i.e., higher likelihood of classifying a novel compound word as real word. Here, we tested two assumptions. First, we predict that schema congruency (congruent versus neutral) should increase error rates for intact versus recombined compound words in a lexical decision task. Second, as we found evidence that schema congruency enables the formation of unitized representations, we hypothesize that schema congruency (congruent versus neutral) should exert a larger influence on associative as compared to item memory in the lexical decision task.

6 Schema-based learning affects performance in implicit memory (Exp. 2)

6.1 Introduction

Implicit or non-declarative memory has been characterized by facilitated task performance without conscious recollection (Graf & Schacter, 1985). In the processing mode model, Henke (2010) defines a non-hippocampal, fast-acting memory processes that operates on rigid, unitized representations which is not restricted to explicit familiarity processes during retrieval but also feeds into non-declarative priming processes. In line with his, Wang and colleagues (2010) state that “[i]mplicit memory effects can emerge as a byproduct of reprocessing the perceptual or conceptual aspects of an item” (p. 835), resulting in respective perceptual or conceptual priming effects.

In using neuroimaging data and data on lesion patients, Wang and colleagues (2010) show that the perirhinal cortex together with other regions in the medial temporal lobe (MTL) underly conceptual implicit memory, measured in exemplar generation and semantic decision tasks. This is of special interest, as the perirhinal cortex has been argued to underly the formation of unitized representations, especially if they are later recognized by means of familiarity (Haskins et al., 2008) and has been considered to underly item recognition (Davachi et al., 2003). In referring to research by Wang and colleagues (2010; 2014), Parks and Yonelinas (2015) summarize that fluency in processing can not only support familiarity-based recognition, but also implicit memory.

In their second experiment, Parks and Yonelinas tested whether unitization influences item and associative memory performance in an implicit lexical decision task (Meyer & Schvaneveldt, 1971). In line with the idea that unitization constitutes a specific learning mechanism dissociable from levels of processing (Craik & Lockhart, 1972), they expected that unitization should preferentially benefit associative memory over item memory performance. The definition-sentence paradigm (e.g., Bader et al., 2010; 2014; Haskins et al., 2008; Quamme et al.,2007; Wiegand et al., 2010) was applied in a within-subjects manner. Participants underwent one block of the definition condition and one block of the sentence condition whereby block order was counterbalanced across participants. Afterwards, participants completed a lexical decision task in which they were shown intact, recombined and yet unpresented novel compound words together with real compound words. Their task was to decide for each compound word whether it was real or fake. The dependent variable was the percentage of incorrect word responses for novel compound words and from this, priming scores were derived in subtracting the error rate on novel, not learned, compound words from the error rate on intact and recombined compound words. The novel, unpresented compound words were highly similar to the learned compound words and thus serve a realistic baseline for processing of novel compound words. As expected, participants did more often erroneously classify intact than recombined word pairs as real words, thus, they showed associative priming effects in the unitization condition. There were no significant associative priming effects in the no-unitization condition. The interaction between stimulus type (intact, recombined) and unitization (yes, no), i.e., the critical test of the idea whether associative priming effects were larger in the unitization than in the no-unitization condition, was marginal significant (p = .051). They did not find a significant item priming effect in the unitization condition. Thus, this experiment provides first evidence in that unitization also benefits associative memory performance in an implicit memory test.

Based on the results of experiment 1 and the second experiment from Parks and Yonelinas (2015), we reasoned that if schema-congruency influences recognition memory by triggering the formation of unitized representations, elevating the contribution of familiarity to recognition memory, then schema-congruency should similarly affect performance in an implicit memory task.

We tested this idea in an additional behavioral experiment, which was conducted online. Here, the incidental learning task used in experiment 1 (chapters 4 and 5) and experiment 3 (chapter 7) was adapted by reducing the number of learned trials to 120. Half of these 120 compound words were learned with a congruent definition preceding the compound word and the other half with a neutral definition. After a retention interval, a lexical decision task followed in which participants had to decide for each presented compound word if it was a real, i.e., already existing German compound word, or not. Next to real compound words, the compound words from the encoding phase were presented together with novel, yet unpresented compound words. If schema-congruency triggers the formation of unitized representations and thereby semantic integration of the compound word (Mecklinger & Bader, 2020), conceptual fluency should be higher in the congruent condition. This elevated fluency in the congruent condition might then boost novel compound words near to the level of real compound words, resulting in more false alarms on novel compound words or higher reaction times for correctly rejected novel compound words.

We formulated two main hypotheses. First (H1), there should be statistically significant priming effects on errors (call a novel compound word an existing word) for intact and recombined compound words in both congruency conditions.

Second (H2), schema congruency (congruent vs. neutral) should modulate associative priming in that there should be larger performance costs (higher Error-priming scores) for intact as compared to recombined compound words in the congruent vs. the neutral condition.

6.2 Methods

6.2.1 Open Science Statement

The current experiment has been pre-registered with AsPredicted (https://aspredicted.org/). The pre-registration can be seen from https://aspredicted.org/pc74x.pdf. Whenever we depart from what has been pre-registered, this is mentioned in the text.[8]

6.2.2 Estimation of required sample size with an a priori power analysis

The minimal required sample size of $N = 75$ was determined with an *a priori* power analysis, based on the effect of interest, i.e., the interaction between Levels of unitization (high, low) and type (intact, rearranged). The effect size was estimated from Parks & Yonelinas (2015), Experiment 3. Cohen's d_z for the t test of the difference between differences was calculated from the reported F value, $d_z = .29$. Sample size estimation was done for a paired samples t test with $\alpha = .05$, $1\text{-}\beta = .80$. To obtain a fully counterbalanced design, we aimed to draw $N = 84$ subjects that pass our inclusion criteria.

6.2.3 Inclusion criteria

We pre-registered several inclusion criteria. Participants were recruited via Prolific (www.prolific.co) [2022]. The prolific filter criteria have been used to draw a sample which constitutes 18–30-year-old German native speakers, who are right-handed and do not report having any ongoing mental health/illness or condition and reported agreement with taking part in deception studies. The latter criterion was applied as the current experiment is an incidental learning experiment in which participants are not instructed that the experiments goal is investigating memory performance.

[8] After having collected the data, we realized that hypothesis 3 cannot be statistically adequately tested with the current design. Therefore, we do not report hypothesis three or related analyses here.

6.2.4 Participants

We collected N = 122 complete datasets. To those, the pre-registered exclusion criteria were applied.

At first, we checked whether recording quality was appropriate for the current study. Here, we followed recommendations from the LabVanced documentation[9] and excluded data sets for which the recorded timing quality value (standard deviation of all time measures) is larger than two to three times the median standard deviation (12 ms) which we rounded up to 40ms. This was the case for n = 20 data sets. Second, datafiles were excluded when participants stated having studied stimuli during the to-be-incidental learning phase, because they expected them to be later asked for in a memory test (n = 18 participants) in a survey at the end of the experiment. Third, as the experiment was conducted in a less controllable online setting, we included a learning criterion to ensure participants attended to the task. Therefore, classification as congruent or neutral during learning should not be statistically different from chance in a binomial test (n = 9). Lastly, we defined that in the lexical decision task, trials with a reaction time under 200 ms and complete data sets consisting of more than 20% of such trials would be excluded. No dataset was excluded because of this criterion. Moreover, data sets would also be excluded if the same response was made in more than 90% of the trials. We additionally pre-registered that datasets would be excluded if less than 90% of the trials were responded to. However, we noticed that this filter criterion is no good choice when the dependent variable of interest is error rates. Therefore, we deviated from the pre-registration in this aspect. No dataset was excluded because of this criterion. We pre-registered an additional criterion, which is that participants would be excluded if they report familiarity with the learning in a survey at the end of the experiment. The question asked was:

„Ist Ihnen im Nachhinein aufgefallen, dass Sie die zusammengesetzten Wörter und Sätze bereits aus einer anderen Studie kennen?“

(Did you notice afterwards that you already know the compound words and sentences from another study).

[9] Obtained from https://www.labvanced.com/faq_eng.html (19.04.2023)

The item was included for the rather unlikely case that participants took part at a related study at Saarland university. However, *n* = 31 out of *N* = 122 participants claimed having known the data from another experiment. As this proportion seemed unreasonably large and unlikely, we refrained from applying this criterion.

Thus, the pre-registered exclusion criteria resulted in the exclusion of *n* = 38 datasets. Another *n* = 4 datasets were excluded, as the datasets included extreme values (Abs(z) > 3.29) that were excluded to improve normal distribution of the data as parametric tests were used for data analysis. The final sample consists of *n* = 80 datasets.

Participants gave their informed consent by clicking on a respective button at the beginning of the experiment and were reimbursed with 10£ (the experiment took approx. 1 hour). Participants were debriefed after the experiment. The experiment was approved by the ethics committee of the Deutsche Gesellschaft für Sprachwissenschaft (#2017-07-180423).

6.2.5 Stimulus materials

The lexical decision task was adapted from Parks & Yonelinas (2015) and constitutes 20 original 'intact' congruent compound words, that were presented as they were shown in the preceding incidental learning phase, 20 neutral intact compound words, 20 congruent recombined compound words 20 neutral recombined compound words, 20 unpresented 'new' compound words and 100 existing 'real' compound words. The real compound words were included to make the task plausible.

Hereby, not only an equal distribution of novel compound word types (intact congruent, intact neutral, recombined congruent, recombined neutral, new) is achieved but also an equal distribution of response categories for novel and real compound words (real word, not real word). The latter is important to avoid response biases that could infer with the task.

To achieve this stimulus distribution in the lexical decision task, the preselected 140 compound words were divided into seven bundles (For more details on stimulus selection, please refer to chapter 3.3).

There were no significant differences in word length, sentence length, definition fit ratings, semantic unrelatedness ratings or imagineability ratings across the bundles and no differences in word length to the real compound words. From these bundles, seven learn and test lists were created. Stimulus counterbalancing across lists was achieved by using a latin-square logic whereby the bundles were numbered from 1 to 7. For the learn lists, compounds from 3 subsequent bundles were presented with their congruent definition and compounds from the next 3 bundles with their neutral definition. The remaining bundle was not used during learning, as these stimuli served as new (unpresented) stimuli for the lexical decision task. This assignment was rotated seven times. Like this, within the experiment, each compound word is used together with its congruent and also with its neutral definition. The resulting test lists were created analogous following a fixed scheme whereby the first bundle of the three congruent ones was presented as intact and the other two[10] with their recombined compound word. The same scheme was applied for the neutral bundles. The remaining bundle served as new compound words.

6.2.6 Procedure

The experiment was programmed and hosted via the online experiment platform LabVanced (Finger et al., 2017), and participants were recruited via Prolific (2022; www.prolific.co [2022]). Participants could take part with either a computer or a laptop, but tablets and mobile phones were not allowed. For technical reasons, browsers were restricted to Google Chrome, Mozilla Firefox and Microsoft Edge.

After participants agreed with the consent form, they were instructed to start the experiment in a silent room and make sure that they will not be interrupted. All stimuli were presented in white font against a black background.

The first part of the experiment constituted the incidental learning task. As no EEG was recorded, the jittered fixation cross before the compound word presentation was set to a fixed time of 1000ms. Participants were instructed to evaluate how well new compound words are explained by preceding definitions.

[10] Note that two original compound words are required for one recombined compound word. Therefore, to achieve a similar distribution of intact and recombined compound words in the lexical decision task, two times as many compound words are required.

They were shown stimulus examples and before the main task, they underwent 12 practice trials to familiarize with the task. Afterwards, participants were shown 120 learning trials in 4 blocks of 30 trials each. Trial selection was pseudo-random with the constraints of not more than 3 consecutive trials from the same condition (congruent, neutral). A trial started with a fixation cross, presented for 500 ms. After, the main sentence was presented for 1000 ms. Afterwards, the relative sentence was added to the screen and the whole definition was presented for another 3500 ms. Afterwards, a fixation cross was shown for 1000 ms followed by the compound word that was presented for 2000 ms. After another 500 ms blank screen, the response screen was shown for up to 3000 ms or until participants gave their response. They should indicate for each compound word how well it was described by the preceding definition by using a 4-point scale (1, not at all; 2, rather not; 3, rather; 4, absolutely) mapped onto the keys x, c, n and m. For half of the participants, key assignment was ascending and for the other half descending. The response screen contained the question participants should answer of how well the compound word was described by the preceding definition and also the response scale with the key assignment they should use. A 500 ms- blank screen followed until the next trial started.

After the incidental learning phase, a retention interval of approximately 10 minutes followed in which participants were presented with correct or incorrect easy calculations and should decide for each one if it was correct or incorrect. The task was chosen to hinder participants from thinking about the previous task and to remove stimuli from working memory.

Afterwards, the lexical decision task followed. They were instructed that they would be presented with compound words and should decide (as fast but correct as possible) for each one if it was an existing word of German (“real”) or not (“unreal”). They were also told that they could respond as soon as the word is presented but before it is not shown anymore. They could familiarize with the task in 20 practice trials. After, they were shown 200 lexical decision trials. Trial selection was pseudo-random with the constraints of not more than 4 consecutive trials from the same type (congruent intact, neutral intact, congruent recombined, neutral recombined, new, real). A trial consisted of a 500 ms- fixation cross, followed by the compound word in the center of the screen, which was presented for up to 1000 ms or until a response was given. Below, the question “Real word” was shown together with the response labels (real or unreal) and the keys. Participants used the keys c and m whereby key assignment to real and unreal was counterbalanced across participants.

The last part consisted of a short survey in which participants were asked if all stimuli (words and sentences) were presented completely on the screen, if they realized that they already knew the compound words and sentences from another experiment and if they tried to memorize the stimuli in the first part of the experiment because they expected a memory test to occur later. After the survey, participants were thanked for their participation, debriefed and re-directed to prolific were they obtained monetary compensation for their time.

6.2.7 Data analysis

For all analyses, the significance criterion of $p < .05$ was applied. Data were analyzed using R (version 4.2.1; R Core Team, 2022) and RStudio (Version 2022.7.2.576; RStudio Team, 2022). The packages *dplyr* (version 1.0.10; Wickham et al., 2022), *tidyverse* (Wickham et al., 2019), *naniar* (version 0.6.1; Tierney et al., 2021), *pwr* (version 1.3-0; Champely, 2020), *pastecs* (version 1.3.21; Grosjean & Ibanez, 2018) and *afex* (version 1.2-0; Singmann et al., 2022) were used. Whenever non-hypothesis-driven multiple testing was required, the Bonferroni-Holm correction (Holm, 1979) was applied. The reported corrected *p*-values were calculated with the function p.adjust of the R package *stats* (R Core Team, 2022). Measured dependent variables are errors (i.e., erroneously classifying a novel compound word as real compound word) and reaction times. Error scores were calculated as the percentage of errors relative to analyzed trials[11] (all trials with a reaction time of > 200ms). Thereof, we derived error priming scores error rates on new compound words (unpresented during learning) from error rates for intact and recombined compound words, respectively. New compound words serve as a baseline for the processing of

[11] We believe this is the most adequate way to account for trials with a reaction time below 200ms, as those trials should be invalid of lexical decisions and thus add noise to the data. However, we analyzed whether this decision has consequences for the results. Changing this trial criterion did not qualitatively change the effects when interactions are interpreted one-sided, licensed by our specific hypotheses.

novel compound words, as they are highly similar in type, but not influenced by the experimental manipulation.

The three hypotheses were then analyzed by means of repeated measure analyses of variances (ANOVAs) and paired-samples *t* tests on erroneous word responses (Parks & Yonelinas, 2015). Behavioral outliers were defined as extreme values, i.e., with a standardized z-value greater than 3.29 above the mean (Field, 2009, p. 179). As measures of effect size, we report Hedges' g_{av} for effects from paired samples t tests with the formula provided in the spreadsheet (Version 5; Lakens, 2013) and partial eta squared (η^2).

6.3 Results

As a manipulation check, we first tested whether congruency ratings in the learning phase were given in accordance with our congruency manipulation. Therefore, we calculated a paired samples *t* test comparing congruency judgements in the congruent and in the neutral condition. As expected, congruent trials were rated as significantly more congruent (i.e., higher ratings) than neutral trials, $t(79) = 39.60$, $p < .001$, $g_{av} = 6.01$ ($M_{congruent} = 3.34$, $SD_{congruent} = 0.28$, $M_{neutral} = 1.64$, $SD_{neutral} = 0.28$).[12]

Table 3 shows the reaction times on correct classifications in the lexical decision task. An exploratory congruency (congruent, neutral) x stimulus type (intact, recombined) – ANOVA on reaction times did not reveal any significant effects (all *ps* > .411).

Table 3

Reaction times

	Schema congruency during learning	
Stimulus type	Congruent	Neutral
Intact	765 (69)	757 (71)
Recombined	764 (66)	750 (78)
	New items (during lexical decision task)	
Novel	755 (66)	
Real	677 (52)	

Table 4 shows the proportion of compound words for which a word response was given in the lexical decision task. Real compound words were only included to make the task plausible but are not further considered in the analyzes.

[12] This test was not pre-registered, but the same procedure was used in all three experiments of this thesis.

Table 4

Error rates

	Schema congruency during learning	
Stimulus type	Congruent	Neutral
Intact	.23 (.16)	.21 (.15)
Recombined	.17 (.14)	.18 (.17)
	New items (during lexical decision task)	
Novel	.14 (.13)	
Real	.13 (.10)	

Note. Standard deviations are shown in parentheses. All stimulus types are considered unreal words in the lexical decision task, except from the real compound words.

We first calculated one sample *t* tests to check whether the obtained error priming scores are larger than zero (H1). Thus was the case for both, the intact and the recombined compound words in both conditions (all $p_s < .001$) meaning that participants made significantly more errors on compound words presented as such or recombined during learning than to novel compound words. Next, we tested the pre-registered hypotheses that there should be larger associative priming effects in the congruent as compared to the neutral condition (H2) and that associative priming effects should be larger than item priming effects in the congruent, as compared to the neutral condition (H3).

6.3.1 Are there larger associative priming effects in the congruent versus the neutral condition?

Priming scores for intact and recombined compound words can be seen from Table 5.

Table 5

Priming scores for intact and recombined compound words

	Schema congruency during learning	
Stimulus type	Congruent	Neutral
Intact	.10 (.10)	.07 (.12)
Recombined	.04 (.09)	.04 (.11)

Note. Standard deviations are shown in parentheses.

To test whether associative priming effects are larger in the congruent as compared to the neutral condition, we calculated a congruency (congruent, neutral) x item type (intact, recombined) – repeated measures ANOVA on error priming scores. We did not find a significant main effect of congruency, $F(1,79) = 1.02$, $p = .316$, $\eta^2 = .01$, but a significant main effect of item type, $F(1,79) = 20.96$, $p < .001$, $\eta^2 = .21$, as well as a significant interaction between congruency and item type, $F(1,79) = 4.03$, $p = .048$, $\eta^2 = .05$.

To resolve the significant interaction between congruency and item type, we calculated follow-up paired-samples *t* tests in which we compared error priming scores for intact and recombined compound words for each condition, separately. Those tests revealed a significant associative priming effect with larger error rates for intact versus recombined compound words in the congruent condition, $t(79) = 4.94$, $p < .001$, $g_{av} = 0.61$(one-sided), and in the neutral condition, $t(79) = 2.43$, $p = .017$, $g_{av} = 0.25$ (two-sided). The descriptive statistics can be seen from Table 5. Thus, there were significant associative priming effects for both the congruent and the neutral condition, but the associative priming effect was larger in the congruent, as compared to the neutral condition, resulting in the significant interaction.

6.3.2 Are there item priming effects in the congruent versus the neutral condition?

As Parks & Yonelinas (2015), we exploratorily tested whether congruency modulated item priming. Associative and item priming scores for congruent and neutral compound words can be seen from Table 6. Item priming scores equal recombined error priming scores.

Table 6

Associative and item priming scores

	Schema congruency during learning	
Priming type	Congruent	Neutral
Item	.04 (.09)	.04 (.11)
Associative	.06 (.11)	.03 (.11)

Note. Standard deviations are shown in parentheses.

We calculated a paired-samples *t* test comparing item priming effects across conditions. There were no significant differences for item priming between the congruent and the neutral condition, $t(79) = -0.47$, $p = .644$, $g_{av} = 0.05$ (two-sided). The descriptive statistics can be seen from Table 6.

6.3.3 Further exploratory analyses

In addition to the results of erroneous word responses on novel compound words, we exploratorily analyzed error rates including trials in which no response was given and also reaction times on incorrect word responses. However, those analyzes did not show any of our expected interaction effects.

6.4 Discussion

The goal of experiment 2 was to test whether the consequences of schema-based learning can also affect performance in an implicit memory task. Therefore, a lexical decision task similar to Parks & Yonelinas (2015) was used, in which novel compound words (original, recombined, new) were presented together with real noun-noun compound words. The idea of this experiment was that if compound words are already integrated in semantic memory, this integration should elicit conceptual fluency (Mecklinger & Bader, 2010) to a comparable level than real compound words. Thus, it should be harder to reject them as not real. In line with the results from Parks and Yonelinas (2015), we tested whether there was larger associative priming in the congruent as compared to the neutral condition, taking account performance of recombined compound words as a more adequate measure of semantic integration of the novel concept as a whole. In line with our hypotheses, we found evidence towards larger associative priming in the congruent as compared to the neutral condition. Of note, our results indicate that in line with the reasoning by Henke (2010), unitized representations can be used in both, explicit and implicit memory tasks, several minutes after initial encoding and independent from the presence of a retrieval intention.

Based on the absolute familiarity effect we found in experiment 1, which is considered to be semantic in nature (see Mecklinger & Bader, 2020), it is conceivable that the higher tendency to make a real word response for a novel compound learned in the congruent condition originates in higher conceptual fluency which is deemed diagnostic for lexical status. To support this idea, it would be interesting to repeat experiment 2 with event-related potentials to test whether false real word responses are accompanied by an N400 attenuation effect similar to experiment 1.

6.5 Conclusion & next steps

In two experiments, we provided evidence in that associative memory performance benefits from schema congruency. However, recent research indicated that learning under the influence of a schema does not always lead to better memory and it was even argued that schematic encoding leads to perceptually less detailed memory representations (Gilboa & Marlatte, 2017), resulting in less discriminability between learned and similar information (e.g., Spalding et al., 2015; Sweeger set al., 2015; Webb et al., 2016; see Gilboa & Marlatte, 2017 for a review). In line with this idea, we assumed that whether schema congruency is beneficial for memory performance or not might depend on the type of associative information that is used as distractors in the memory test: If, as a result of schema-congruency, two items are integrated into a unitized representation, that is probably semantic in nature (indicated by the N400 absolute familiarity effect), it might be harder to distinguish original from semantically similar information. In contrast, distinguishing intact from recombined compound words might be supported by a less precise semantic representation, as recombined compound words violate the concept of the compound word in which the constituents were originally used.

In experiment 3, we tested whether schema-based learning of novel word pairs leads to an imprecise memory representation, as a result of which it is harder to distinguish original from highly similar information. Alternatively, unitized representations as new entities could be protected against false memory effects (see Tibon et al., 2018).

7 How precise are memory representations formed under the influence of a schema (Exp. 3)?

7.1 Introduction

Was the novel compound word you read about in the newspaper *flight shame* or *flight guilt*? If the word *flight shame* has been learned under the influence of a schema, it is likely that you erroneously believe it could have been *flight guilt*, which also matches the concept. This idea of overusing a schema (see Bartlett, 1932) has been referred to as schema generalization. However, rather heterogeneous phenomena can be subsumed under this umbrella term. Recent research indicated that schematic encoding leads to perceptually less detailed memory representations (Gilboa & Marlatte, 2017), resulting in less discriminability between learned and similar information (e.g., Spalding et al., 2015; Sweegers et al., 2015; Webb et al., 2016; see Gilboa & Marlatte, 2017, for a review).

Spalding and colleagues (2015) presented participants with congruent or incongruent pairings between object pictures and context words in a congruency judgement task. In this task, participants should visualize the objects in the contexts given by the word and rate the congruency of this relationship on a four-point scale. During a subsequent recognition test phase, participants had to discriminate learned pictures from perceptually similar foils in deciding for each object if it was old, similar, or new. Two groups of participants took part in the experiment, a healthy control group and a patient group with lesions in the ventromedial prefrontal cortex (vmPFC), a brain area which is thought to be crucial for schema-based learning (see e.g., van Kesteren et al., 2012). The patient group perceived congruency (indicated by rating values) different to the control group, agreeing less with the congruent-incongruent trial classification than the control group, supporting the idea that the vmPFC might play a role in schema processing. Recognition performance on old and new trials did not differ across groups, showing that the patients did not show altered recognition performance in general. However, the healthy control group showed a congruency effect in that they were more likely to classify old and similar objects as old when they had been presented in a congruent than an incongruent pairing during learning. The patient group did not show a congruency effect, even if the statistical model accounted for the altered congruency ratings during learning. The pattern of results indicates that the beneficial effects of schema congruency for memory performance do not show for highly similar information, probably because if a less precise memory representation being formed under the influence of a congruent schema.

Webb and colleagues (2016) investigated memory for originally learned *target objects* and conceptually related or unrelated *lure objects* from before learned visual scenes. During encoding, participants were presented with visual schematic scenes, e.g., of a Bathroom. Their task was to learn the scenes for a later memory test. In the memory test, target objects from the scene as well as yet unpresented conceptually related and unrelated lure objects were shown. Their task was to classify each object as ‘Remember/Know/New’, whereby *remember* meant that they could recollect specific details about the object as shape, color or location within the scene or internal thoughts during learning. *Know* meant that they perceive the object as being familiar without being able to recollect details. *New* meant that the picture was not presented during learning. The behavioral results show higher false alarm rates for both, remember (recollection) responses and know (familiarity) responses in the schema-congruent condition as compared to the incongruent condition. Thus, in two experiments using pictorial materials, schema congruency has been found to benefit memory for originally learned information at the expense of memory performance on similar information. However, those experiments differ from the current work in several aspects. First, in the experiments by Webb et al., (2016) and Spalding et al. (2015), pictures have been used as memoranda, whereby the current work deals with learning of compound words. Second, and related to this, those experiments tested memory performance for pre-existing semantic knowledge whereby in the current work, we are interested in learning novel concepts based on schematic knowledge of existing concepts. Thus, it remains unclear whether unitized representations, formed under the influence of a schema are prone to errors on semantically similar information.

In a series of experiments by Tibon and colleagues (2018), they tested a related idea, namely if by means of a congruent definition, a schema representation is acquired that can be rapidly used and support learning of another schema congruent word pair (compound word) that are harder to correctly reject in a later recognition memory test. In an initial learning phase, novel word pairs (e.g., CLOUD LAWN) are learned together with a congruent definition, enabling unitized learning, or a sentence, not enabling unitized learning (referred to as the definition-sentence paradigm in the following). Participant's task was to judge the fit of the word pair with either the definition or the sentence. A relearning phase follows, in which the same word pairs are presented again (CLOUD LAWN; "repeat"), the first word is substituted by a semantically related word (MOON LAWN; "related") or a semantically unrelated word (TEA LAWN; "unrelated"). Here, participants are instructed to judge which of both words is more frequent in their daily life. In the later test phase, a 4-alternative forced choice test was used; Participants were presented with the second word of a word pair together with a word that was old-different (i.e., the first word in another word pair), new-related (i.e., not presented but semantically related to the presented word) and new-unrelated (i.e., yet unpresented) and had to select the correct pair-associate from relearning. The authors assumed that the unitized representation can be thought of as a schema, participants should show a pattern of schema generalization, i.e., they should make more false alarms on related words in the unitization condition than in the non-unitization condition. They did not find evidence in favor of such a schema account of unitization in three experiments. Of note, in their relearning design, word pairs were introduced as arbitrary word pairs which might have hampered context-dependent schema instantiation. However, as the presentation of the word pair is still highly similar to the original presentation, this might constitutes a valid manipulation, and this experiment provides some first support for this idea that the schema representation including conceptual knowledge might not be rapidly viable. Thus, we tentatively conclude that this data speaks against the idea that a unitized representation can be thought of as a schema representation.

How can the results by Tibon and colleagues (2018) be explained in the context of the results by Webb and colleagues (2016) and Spalding et al. (2015)?

As already elaborated on above, schema-based learning in a sense that existing information is remembered to be part of an experimental episode as in Webb et al. (2016) and Spalding et al. (2015) differs from learning a novel concept as in Tibon et al. (2018) and our work. In the paradigms used in our work and in Tibon et al. (2018), the schema can be thought of as an on-the-spot activation of an already known word-formation rule explaining how the underlying (pre-existing) concepts must be combined to the whole word meaning, i.e., the novel concept. One way to explain the results by Tibon and colleagues (2018) could be to assume that although the memory representation is formed under the influence of a congruent schema, it must not be thought of as a schema itself.

Gilboa & Moscovitch (2021) stressed that encoding of information into memory comprises the formation of several parallel representations. Among those are detailed episodic representations containing event-specific information, a schematic representation holding shared information across several episodes (Gilboa & Marlatte, 2017) and gist representations of general conceptual information underlying an event. For the compound learning case, the episodic representation might constitute the word form *Vegetablebible* together with the definition acquired during learning (*A dictionary used by gardeners*) as well as spatiotemporal context information about the learning situation (e.g., *On Monday, I was at this experiment at Saarland University*). The schema representation might constitute the abstract concept formation rule, provided by the definition, recruiting prior concept knowledge of the word constituents. The gist-like representation, however, might contain a more general, semanticized version of the compound word, including broader concept knowledge.

Theoretically, both the schema representation and the gist-like representation could be the origin of conceptual fluency, the contribution of which we showed in experiment 1 and 2. However, for the following reasons, we believe this is less likely the case for the schema representation than the gist representation. First, if the schema representation in the current experiment is the pre-existing repetitive word formation rule (as Ghosh & Gilboa 2014 would probably agree on) without conceptual knowledge of the underlying word constituents, it would not be specific enough to be diagnostic during the recognition memory test, as it applies to many presented novel compound words. Here, the schema would be the abstract learning rule of how conceptual knowledge must be combined to form a new concept without including underlying conceptual knowledge: The second constituent is the head of the compound word which establishes the main concept. This concept is modified by the concept held by the first constituent in applying the relationship provided by the definition. Those rules are part of our language knowledge of how novel concepts can be formed (Gagné & Spalding, 2009) and thus can be seen as already existing schemata. However, next to probably not being useful during recognition judgements, this highly general schema would theoretically be congruent to compound words in all contexts. Alternatively, if the schema representation should be seen as this word formation rule including the underlying concepts, the schema itself would be new to the system and thus would require strengthening over many repetitions to be stored and thus viable for reactivation (see Cockroft et al., 2022; Gilboa & Marlatte, 2017). Thus, the more parsimonious assumption would be that whilst a schema can form on the spot and support learning (see Hebscher et al., 2019), this does not necessarily mean that the schema itself is permanently stored and can be reactivated after one learning instance. This would explain the results of Tibon and colleagues (2018) in that the acquired representation cannot (yet) be used to learn something new, as it would be expected for a schema representation.

Thus, it is conceivable that schema-based learning leads to the formation of a less-detailed gist-like representation that is not directly (“online”) available during relearning but can be erroneously applied to similar information, explaining false memory effects.

To test this hypothesis, we repeated experiment 1 with slight modifications. Again, novel compound words are presented in an incidental learning task, either preceded by a semantically congruent definition, explaining the combination of the words to the novel concept (schema-congruent condition), or a neutral context that does only explain the contribution of one word constituent to the novel concept (neutral condition).

In the subsequent memory test, semantic lure compound words were presented instead of new, newer seen compound words and should classify compound words as old (have been presented exactly as during learning, only correct for intact compound words) or new (correct for recombined compound words and semantically related lures).

Semantic lures were created by exchanging the second constituent, i.e., the head of the compound word for a semantically related new word. Crucially, we modified the second and not the first constituent of the compound word to make sure that the differential priming of the first constituent introduced by the congruent and the neutral condition does not affect our results. If we find elevated false alarm rates for compound words learned in a congruent as compared to a neutral condition in exchanging the head of the compound word, which is identical in both conditions, this will provide evidence for differences in associative conceptual information underlying our experimental effects. If schema-based processing contributes to the learning of novel compound words, participants should do more false alarms on semantically related lures in the schema-congruent condition as compared to the neutral condition.

7.2 Methods

7.2.1 Open Science Statement

The current experiment has been pre-registered with AsPredicted (https://aspredicted.org/). The pre-registration can be seen from https://aspredicted.org/qj27u.pdf. Whenever we depart from what has been pre-registered, this is mentioned in the text.

7.2.2 Estimation of required sample size with an *a priori* power analysis

The minimal required sample size of $N = 16$ was determined with an *a priori* power analysis, based on the smallest effect of interest, i.e., partial eta squared $\eta^2 = .21$ from the congruency (congruent, incongruent) x item type (old item, lure) interaction on the proportion of old responses from Packard et al. (2017, Exp. 2). The within-subjects MANOVA was initially chosen as analysis type to avoid issues with violated sphericity[13]. The parameters entered in the analysis were $f = 0.52$, $\alpha = .05$, $1\text{-}\beta = .80$, number of groups = 2, number of measures = 4, correlation among repeated measures = 0. To obtain enough observations for a reliable measurement and achieve a fully counterbalanced design, we aimed to draw a sample of $N = 40$ participants.

Unfortunately, after publishing the pre-registration and having collected the data, we became aware that G*Power cannot reliably estimate the required sample size of ANOVAs (Giner-Sorolla et al., 2020), with probably similar issues for MANOVAs. As power is a pre-data concept and thus, achieved power cannot be estimated (Wagenmakers et al., 2015), we choose a different approach to determine whether our drawn sample size would be adequate. We calculated an additional sample size estimation analysis, by changing the analysis type to a paired samples *t* test, to determine the required sample size to detect our effect of interest. Therefore, the *F* value of the interaction from the congruency (congruent, incongruent) x item type (old item, lure) interaction on the proportion of old responses from Packard et al. (2017, Exp. 2) was transformed

---

[13] As no analyses with more than 2 levels per factor are planned, we changed analysis type to ANOVA.

to the respective *t* value. This *t* value refers to the test whether the differences differ. From this *t* value, $d_z = 0.50$ was obtained and entered in a power analysis for a paired samples *t* test with the parameters $\alpha = .05$, $1\text{-}\beta = .80$, two-sided. The R package *pwr* was used. The required sample size, based on this analysis is $N = 34$ (rounded up to the next whole number). Thus, the obtained sample size of $N = 40$ would theoretically result in a sufficiently powered study.

7.2.3 Participants

$N = 48$[14] young adults volunteered for this study, having been recruited via flyers and local databases.

We pre-experimentally determined that datasets would be excluded if the same response were made in more than 90% of the trials, if less than 90% of the trials were responded to or if more than 20% of trials are responded to within less than 200 ms. Three datasets were excluded as they did not respond to at least 90% of experimental trials (learning phase or test phase were considered separately).

Data from $n = 2$ participants had to be excluded because they reported that they intentionally studied the stimuli or took part in an experiment using the same stimuli ($n = 8$). The final sample consisted of $N = 35$ participants (24 females, with an age range from 18 to 30, $Mdn = 22$ years, $SD = 2.90$). All participants were students of Saarland University or volunteers from the community and reported being in good health, not suffering from any neurological or psychiatric conditions and having normal or corrected-to-normal vision. Further, all participants were right-handed, as assessed with the Oldfield Handedness Inventory (Oldfield, 1971), and reported being native speakers of German. Participants gave their informed consent and were reimbursed with 10E/h or course credit. Participants were debriefed after the experiment.

[14] Exclusion of datasets due to prior participation or intentional learning was monitored during data collection and were replaced as far as possible to achieve a counterbalanced design. By this, one additional dataset was collected too much by mistake. As the inclusion of this dataset would have introduced an imbalance in the stimulus lists, it was discarded before the data were analyzed. The additional exclusion of datasets due to statistical criteria introduced a slight imbalance in stimulus lists. However, inclusion of this dataset would have increased this imbalance.

7.2.4 Stimulus materials

The 240 stimuli from experiment 1 were used in the current experiment within identical encoding lists (see chapter 4). Which encoding list was used was fully counterbalanced across participants. To create lists for the test phase, stimuli were further divided into five subsets of 24 compound words, each, by dividing Set 1 and Set 2 in five packages, respectively. This enabled us to counterbalance the type of presentation of a learned compound word in the test phase, i.e., as intact, recombined or lure compound word, by simultaneously achieving a stimulus distribution requiring an equal amount of old and new responses, i.e., 48 intact compound words, 24 lure compound words and 24 recombined compound words per condition[15]. Which form of the compound word (intact, lure, recombined) was presented during the test phase was fully counterbalanced by using a Latin-square design, so that each compound word was equally often presented in its lure form and twice as many times in its intact and recombined form. Thus, for each encoding list, the same 5 test lists were created, resulting in 10 possible combinations of encoding and test lists. Each test list consisted of 96 intact compound words, 48 recombined compound words, and 48 lure compound words. Stimulus presentation in the experiment was pseudo-randomized for the encoding and test phase, with the limitation of not more than 3 consecutive trials in the same context condition (encoding phase) or not more than 3 consecutive trials requiring the same response (test phase). Due to our exclusion criteria, we did not achieve a fully counterbalanced design. However, all stimulus lists were shown approximately equally often.

7.2.5 Procedure

After having given their written–informed consent, participants completed a language questionnaire and the Oldfield Handedness Inventory (Oldfield, 1971). The experiment was created using E Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). The experiment proper consisted of an incidental encoding phase, a retention interval with a duration of about 10 min

[15] As only a single recombined compound word was created out of two intact compound words, twice as many learnt compound words are required for later recombined compound words as for later intact compound words.

and a test phase. All stimuli of encoding and test phase were presented in white font against a black background. The learning phase was identical to experiment 1 with one exception: The jittered fixation crosses before the compound word presentation (in learning and test phase) was set to a fixed time of 1000ms. During incidental learning, participants were presented with 240 definitions, half of them congruent and half of them neutral, followed by the respective novel compound word. Participants were instructed to rate on a scale from 1 (not at all) to 4 (absolutely) how well the novel compound word denotes the concept given by the definition (see chapter 6.2.6 for timing of the learning task). Their task was to rate how well the compound word is described by the preceding definition, providing a measure of the semantic congruency between the definition and the compound word. Participants responded on a keyboard by using the keys ×, c, n, and m with their index and middle fingers of each hand. The scale was ascending for a part of the participants and descending for the other part of participants. Before the encoding phase, participants completed 8 practice trials to familiarize with the task. The encoding phase was followed by a retention interval. During this interval, participants performed two distractor tasks. At first, an adapted computerized version of the Digit Symbol Task (Wechsler, 1955) from Häuser et al. (2019) was performed for approximately 5 min, followed by 3 min of solving mathematical equations. Only then, participants were told about the upcoming test phase. During the test phase, participants were presented with one of five test list versions, consisting of 96 intact compound words, 48 recombined compound words and 48 lure compound words. A trial started with a fixation cross (1000 ms). Then, the compound word was presented (for up to 3000 ms), until participants gave their response. Participants gave their answer on a keyboard by using the keys f and j to indicate if the compound word was old or new. Key assignment was counterbalanced, ensuring that across participants, each response option was used with an equal frequency. After a 500 ms blank screen, participants were asked to indicate their confidence on the previous response (sure or unsure) using their index fingers, whereby key assignment was ascending for a part of the participants and descending for the other part of participants. The confidence scale remained on the screen for up to 3000 ms or until participants gave their response and was only presented if a response had been logged on the compound word. A trial ended with a blank screen, which was presented for 500 ms. In both, learning and test phase, there were self-paced breaks after 60 trials (encoding phase) or 48 trials (test phase), respectively. Before the test phase, there were 3 practice trials during which participants could familiarize themselves with the task.

7.2.6 Data analysis

For all analyses, the significance criterion of $p < .05$ was applied. Data were analyzed using R (version 4.2.1; R Core Team, 2022) and RStudio (Version 2022.7.2.576; RStudio Team, 2022). The packages *dplyr* (version 1.0.10; Wickham et al., 2022), *tidyverse* (Wickham et al., 2019), *naniar* (version 0.6.1; Tierney et al., 2021), *pwr* (version 1.3-0; Champely, 2020), *pastecs* (version 1.3.21; Grosjean & Ibanez, 2018) and *afex* (version 1.2-0; Singmann et al., 2022) were used. Whenever non-hypothesis-driven multiple testing was required, the Bonferroni-Holm correction (Holm, 1979) was applied. The reported corrected *p*-values were calculated with the function p.adjust of the R package *stats* (R Core Team, 2022). To capture associative memory performance by considering intact and recombined compound words irrespective of correct rejections of new compound words, an associative *Pr* (hits - false alarms) was calculated. Therefore, the associative hit rate was calculated as the amount of compound words, correctly identified as old, divided by the sum of all intact trials, classified as either old or new. The associative false alarm rate was calculated as the sum of all recombined items classified as old, divided by the number of recombined items either classified as old or new.

Behavioral outliers were defined as extreme values, i.e., with a standardized z-value greater than 3.29 above the mean (Field, 2009, p. 179). Hypotheses are tested with ANOVAs and paired-samples *t* tests. As measures of effect size, we report Hedges' g_{av} for effects from paired samples t tests with the formula provided in the spreadsheet (Version 5; Lakens, 2013) and partial eta squared (η^2) for results of the ANOVA.

7.3 Results

7.3.1 Perceived schema congruency in the learning phase

As a manipulation check, we tested whether congruency ratings in the learning phase were given in accordance with our congruency manipulation. Therefore, we calculated a paired samples *t* test comparing congruency judgements in the congruent and in the neutral condition. As expected, congruent trials were rated as significantly more congruent (i.e., higher ratings) than neutral trials, $t(34) = 24.03$, $p < .001$, $g_{av} = 4.67$ ($M_{congruent} = 3.11$, $SD_{congruent} = 0.36$, $M_{neutral} = 1.59$, $SD_{neutral} = 0.27$).[16]

7.3.2 Associative memory performance considering intact and recombined compound words

In a first step, we were interested in conceptually replicating the pattern of results from experiment one in that associative memory performance (indicated by an associative *Pr* score) benefits from schema congruency. In experiment 1, this effect could be traced back to the associative hit rates whereby no differences were found for associative false alarm rates.

A paired samples *t* test on associative *Pr* scores revealed better associative memory performance in the congruent than in the neutral condition, $t(34) = 4.88$, $p < .001$, $g_{av} = 0.61$ ($M_{congruent} = .47$, $SD_{congruent} = .23$, $M_{neutral} = .33$, $SD_{neutral} = .23$).

A Congruency (congruent, neutral) x Type (Hit, False alarm) – ANOVA revealed a significant main effect of congruency, $F(1, 34) = 12.70$, $p = .001$, $\eta^2 = .27$, a significant main effect of type, $F(1, 34) = 124.81$, $p < .001$, $\eta^2 = .79$ and a significant two-way interaction, $F(1, 34) = 23.81$, $p < .001$, $\eta^2 = .41$. Follow-up paired samples *t* tests revealed that associative hit rates were higher in the congruent than in the neutral condition, $t(34) = 6.14$, $p < .001$, $g_{av} = 1.02$ (one-sided; $M_{congruent} = .76$, $SD_{congruent} = .11$, $M_{neutral} = .64$, $SD_{neutral} = .13$). Similar to experiment 1, we did not find significant across-condition differences in

[16] This test was not pre-registered, but the same procedure was used in all three experiments of this thesis.

associative false alarm rates, $t(34) = -0.76$, $p = .454$, $g_{av} = 0.09$ (two-sided; $M_{congruent} = .29$, $SD_{congruent} = .19$, $M_{neutral} = .31$, $SD_{neutral} = .17$).

7.3.3 False alarm rates on recombined and lure compound words

The descriptive statistics can be seen from Table 7. Our main interest was in false alarm rates on lure compound words. We expected larger false alarm rates on lure compound words in the congruent than in the neutral condition. This hypothesis was tested with a paired-samples *t* test on percentage of false alarms on lure compound words. This test revealed the expected pattern of results with more false alarms in the congruent than in the neutral condition, $t(34) = 2.33$, $p = .013$, $g_{av} = 0.34$ (one-sided).

The pattern of results indicates that we can find statistically reliable differences in false alarms on semantically related lure compound words, but not on associative false alarms considering recombined compound words. However, the comparison is not direct when associative false alarms are calculated as in our formula. Therefore, we exploratorily tested whether false alarms on recombined compound words (percentage of recombined compound words erroneously classified as old) differ across conditions. In addition, we tested whether across-condition differences are larger in false alarms on semantically related lure compounds than in false alarms on recombined compound words in a Congruency (congruent, neutral) x Type (lure, recombined) – ANOVA on false alarm percentage scores.

Table 7

False alarm rates (percentage values)

	Schema congruency during learning	
Stimulus type	Congruent	Neutral
Recombined	.28 (.19)	.30 (.17)
Lure	.33 (.15)	.28 (.16)

Note. Standard deviations are shown in parentheses.

The Congruency x Type - ANOVA revealed no significant main effect of congruency, $F(1, 34) = 1.34$, $p = .255$, $\eta^2 = .04$, no significant main effect of Type, $F(1, 34) = 0.47$, $p = .496$, $\eta^2 = .01$, but a significant interaction, $F(1, 34) = 5.43$, $p = .026$, $\eta^2 = .14$.

In contrast to the significant across-condition difference in false alarms on semantically related lures (see above), we did not find statistically reliable differences in false alarms on recombined compound words, $t(34) = -0.73$, $p = .472$, $g_{av} = 0.08$.

7.3.4 Reaction times

We did not have any expectations regarding the reaction times on correct responses in the current experiment (see Table 8).

Table 8

Reaction times

	Schema congruency during learning	
Stimulus type	Congruent	Neutral
Intact	1559 (244)	1603 (265)
Recombined	1802 (287)	1757 (267)
Lure	1742 (293)	1708 (264)

Note. Reaction times (in ms; standard deviations are shown in parentheses).

An exploratory Congruency (congruent, neutral) x Type (Intact, Recombined, Lure) -ANOVA on reaction times of correct responses (old for intact, new for recombined and lure) was calculated. Greenhouse-Geisser-correction was applied. This analysis revealed no significant main effect of congruency, $F(1, 34) = 0.65$, $p = .426$, $\eta^2 = .02$, but a significant main effect of Type, $F(1.66, 56.48) = 30.12$, $p < .001$, $\eta^2 = .47$, as well as a significant interaction, $F(1.94, 65.79) = 5.40$, $p = .007$, $\eta^2 = .14$. Bonferroni-Holm-corrected paired-samples t tests showed no significant effects (all $p_s > .151$).

7.4 Discussion

To investigate the hypothesis that schema-based learning leads to the formation of less detailed memory representations, we repeated experiment 1 with slight modifications. In the memory test, semantic lure compound words were presented instead of new, newer seen compound words. Semantic lures were created by exchanging the second constituent, i.e., the head of the compound word for a semantically related new word. Crucially, we modified the second and not the first constituent of the compound word to ensure that the differential priming of the first constituent introduced by the congruent and the neutral condition does not affect our results. If we then find elevated false alarm rates for compound words learned in a congruent as compared to a neutral condition in exchanging the head of the compound word, which is identical in both conditions, this will provide evidence for differences in associative conceptual information, underlying our experimental effects. Third, the response format was changed to old-new whilst still the importance of associative information was highlighted in the instructions.

In this experiment, we could provide a conceptual replication for the behavioral results of experiment 1 in that schema-based learning supports associative memory by means of increased associative hit rates by no across-condition differences in associative false alarms. In addition, we found higher false alarm rates for compound words in the congruent as compared to the neutral condition. The pattern of no differences in false alarm rates on recombined compound words is of especial relevance as if participants had stronger relied on item memory for the underlying constituents in one condition, there should have been more false alarms on recombined compound words in this condition, as compared to the other. As this is not the case, differences in item memory cannot explain the pattern of more false alarm rates on semantically similar compound words in the congruent than in the neutral condition. Thus, this pattern of results indicates that in the congruent condition, a less detailed memory representation is used which must be semantic in nature as it is not sensitive to identical compound word constituents as it is the case for recombined compound words, but rather to the semantic content of the new concept.

Thus, from the results of experiment 3, we can conclude that schema-based learning leads to rapid semantic integration and thereby the formation of a less detailed memory representation. In the framework of Gilboa & Moscovitch (2021), the encoding of information into memory goes hand in hand with the formation of parallel memory representations. Among those are detailed episodic representations containing event-specific information, a schematic representation holding shared information across several episodes (Gilboa & Marlatte, 2017) and gist representations of general conceptual information underlying an event. Hereby, the gist-like representation might be the result of semantic integration, induced by the schema and enable conceptual fluency and thereby familiarity (see Experiment 1).

Based on the neurocognitive model on recognition memory put forward by Mecklinger & Bader (2020), we assumed that semantic integration of the compound word constituents to a novel concept might enable conceptual fluency, resulting in absolute familiarity which is deemed diagnostic for prior occurrence. It is tempting to speculate that the gist-like representation we argued to underly the false alarm effect in experiment 3 is erroneously applied to the similar compound word within an absolute familiarity process by eliciting conceptual fluency. However, based on the design of the current experiment, we cannot determine whether false alarms on semantically similar information are driven by familiarity or recollection.

To test whether the imprecise gist-like representation formed under the influence of a schema and erroneously applied to semantically similar information is recruited within a familiarity or a recollection process, a follow-up experiment would be required in which behavioral measures as the remember/know-procedure (Gardiner, 1988; Tulving, 1985) are combined with ERP measures of familiarity and recollection (see Experiment 1) to explore the neurocognitive processes underlying false associative memory of semantically similar information. Hereby, the remember/know paradigm is the method of choice as it is extensively used in false memory research, enabling comparisons across studies and is applicable on a trial-by-trial level (in contrast to the ROC procedure and the speeded/nonspeeded paradigm). In addition, to strengthen the point supported by the results by Tibon and colleagues (2019), it might be fruitful to conceptually replicate their results in using the congruent-neutral-definition paradigm providing a neater operationalization of schema knowledge. In addition, German compound words (in contrast to English compound words) are typically spelled without a space between the constituents. Thereby, German compound words would automatically be processed as such during relearning (in both conditions). If the schema was not instantiated as it is context-specific and not enough context was provided (i.e., morphological information that the word pair should be processed as a compound word), then there should be more false alarms in the congruent versus the neutral condition after relearning. If the schema is indeed not rapidly viable to learn new information, false alarm rates should not differ.

7.5 Conclusion & next steps

In this experiment, participants incidentally learned novel compound words, including a novel concept in either a congruent condition, in which a schema was established, or in a schema neutral condition. In a later recognition memory test, participants showed higher associative hits in the congruent as compared to the neutral condition whilst associative false alarm rates did not differ (including performance on intact and recombined compound words). Strikingly, false alarm rates on semantically similar compound words, for which the second constituent was exchanged were higher in the congruent as compared to the neutral condition. We interpret this pattern of results in that schema-based learning leads to the formation of a less detailed memory representation which is erroneously recruited in recognition judgements leading to higher false alarm rates on semantically similar information.

8 General discussion

Within the current work, we were interested in shedding light on a rather old topic of psychological memory research- the question of how prior knowledge influences the acquisition of new memories. Schema-based learning is woven through the memory psychological literature, dating back to at least the 1930s when Bartlett (1932) formulated his popular theory of *Remembering*. Now, 90 years later, this big question is still far from being exhaustively resolved. However, some promising steps towards an explanation of prior knowledge-based learning have been done. In recent years, mainly initiated by the seminal studies from Tse and colleagues (2007; 2011), the focus of the schema concept has shifted towards a neuronal, brain-systems oriented approach with particular focus on the question how prior knowledge informs learning (Gilboa & Marlatte, 2017; Gosh & Gilboa, 2014; van Kesteren et al., 2012). In an attempt to add to current neuroscientific approaches (e.g., Hebscher et al., 2019; Preston & Eichenbaum, 2013; van Kesteren et al., 2012), in the current work, we adopted a functional perspective on schema-based learning of novel associations.

Neuroscientific research put forward the idea that schema-based learning does not rely on the hippocampus (van Kesteren et al., 2012) or at least recruits the hippocampus less (Tse et al., 2007; 2011) than traditional associative learning. Based on this, we hypothesized that schema-based learning might constitute a specific learning mechanism. Based on a processing mode model on associative learning, we assumed that schema-based learning of novel associations, i.e., novel compound words, might enable the formation of unitized representations by means of rapid semantic integration (see Mecklinger & Bader, 2020). To address this idea, we formulated three research questions, which are: First, what are the neurocognitive mechanisms underlying schema-based encoding. Second, what type of memory representation is formed and what neurocognitive processes underly their use in recognition and third, how can these representations be retrieved.

8.1 Summary of Results of the three experiments

In a series of experiments, we explored schema-based learning of novel word associations, i.e., novel compound words, forming a new concept. Learning was incidental, which we believe is beneficial to show that schema-based learning enables the use of a specific learning mechanism which, however, could be overshadowed by the dominance of the traditional episodic associative learning mechanism (see e.g., Tompary et al., 2020).

Schema congruency was established by means of a congruent definition, comprising conceptual knowledge and a higher-order abstract word formation rule. In the neutral (control) condition, the underlying schema did not include the first constituent of the compound word, thereby not enabling schema-based associative learning of the compound word. All three experiments included an incidental learning phase with highly similar trial procedures, in which novel compound words were learned together with a preceding definition. The definition was congruent for half of the compound words and neutral for the other half of the compound words and participants should evaluate how well the compound word can be explained by the preceding definition on a four-point scale. After a retention interval of about 10 minutes, a surprise memory test followed. This test was an explicit recognition memory test including intact and recombined compound words as well as new, yet unpresented compound words (experiment 1 and 3) or semantically similar lure compound words (experiment 3). Participants were instructed to classify each compound word as intact, recombined, or new (Experiment 1), as old (intact) or new (recombined, similar lures; Experiment 3) or performed an implicit memory task in which they should make lexical decisions (Experiment 2).

8.1.1 What are the neurocognitive mechanisms underlying schema-based encoding?

In Experiment 1, we were interested in the neurocognitive processes underlying schema-based encoding. Here, we found evidence for three processes underlying schema-based encoding, which are semantic priming, establishing schema congruency, indicated by an N400 attenuation effect, the formation of a conceptual compound representations, indicated by a parietal subsequent memory effect that has been argued to reflect the formation of unitized representations (Kamp et al., 2017), and bottom-up semantic integration of the compound word constituents, indicated by an N400 subsequent memory effect. Crucially, only the formation of the conceptual compound representation differs across conditions and can account for schema-based learning. Similar to the recognition memory ERPs, ERPs were delayed, also including the N400 which might be for similar reasons as outlined before.

8.1.2 What type of memory representations is formed and by means of which neurocognitive processes are those representations used during recognition?

Experiment 1 showed that schema-based learning indeed leads to better associative memory performance in the congruent as compared to the neutral condition. This memory advantage can be traced back to higher associative hit rates in the congruent as compared to the neutral condition whilst associative false alarm rates did not differ across conditions. This pattern of results was replicated in Experiment 3 with a slightly different recognition memory test phase, including semantically related lure compound words instead of new, yet unpresented compound words and an old/new instead of an intact/recombined/new classification. Therefore, this pattern of results can be seen as robust and shows that schema-based learning leads to better associative memory when intact and recombined compound words are considered. Crucially, schema-congruency is not always beneficial for memory performance as participants make more false alarms on semantically similar compound words in the congruent versus the neutral condition (Experiment 3). Thus, if memory for associative information is not tested by recombining items, which means that the gist structure is completely violated, but with semantically similar compound words, schema-based learning leads to worse memory performance, probably due to relying on a less detailed gist-like representation. In Experiment 1, we additionally found an interesting data pattern which is that associative memory performance (operationalized with an associative *Pr* score) benefitted more from schema congruency than item memory performance (operationalized with an item memory *Pr* score). This pattern of results has been argued to indicate unitization (Parks & Yonelinas, 2015) and thus is in line with our idea that schema-based learning of novel associations might support the formation of unitized representations.

These behavioral results were supported by results from event-related potentials. In line with a unitization account of schema congruency, we found a larger early, broadly distributed old/new effect which we argue constitutes an N400 attenuation effect (see also Bader et al., 2010; Mecklinger & Bader, 2020). In line with the account put forward by my colleagues Axel Mecklinger and Regine Bader, we interpreted this effect to reflect semantic integration which occurred under the support of a schema, resulting in a unitized memory representation which leads to conceptual fluency and thus absolute familiarity. The late parietal old/new effect, the putative correlate of recollection, showed a numerical (statistically not significant) trend towards being larger in the congruent as compared to the neutral condition. This pattern of results was quite surprising when considering the similar study by Bader and colleagues (2010) where the congruent definition condition was accompanied by less recollection. As this effect was not statistically reliable in the current experiment 1, we tentatively interpreted the difference across the studies in that the detailed definitions might have led to more detailed memory representations in the current study (Craik & Lockhart, 1972), elevating the contribution of recollection to recognition memory decisions.

Of note, compared to prior studies, those ERP effects were shifted in time and occurred later in the current study. This delay was explained with differences in presentation format of the compound words (without a space in the current experiment). As the spatial distributions of the effects are highly similar to the distributions of the effects of prior studies (Bader et al., 2010) and more importantly, the spatial distributions of the earlier and the later effect differed statistically significant, we tentatively concluded that this pattern of results provides first evidence in that schema-based learning supports the formation of unitized representations that can be used by means of absolute familiarity in a recognition memory test. This conclusion is also supported by the behavioral data pattern of more benefits for associative as compared to item memory. Additional empirical support for semantic integration or the formation of a gist-like representation comes from the result of more false alarms on semantically similar compound words in the congruent versus the neutral condition, which must be driven by associative similarity (Experiment 3) and from the result that schema-congruency leads to associative priming in an implicit lexical decision task (Experiment 2).

8.1.3 How can these representations be retrieved?

In experiment 2, we tested whether schema-based learning can also affect memory performance in an implicit lexical decision task. This research question was motivated by the prediction by Henke (2010) in that representations formed under the influence of a unitization process should be accessible in both, implicit (priming) and explicit (familiarity) memory tests. First support for this idea came from the second experiment in Parks and Yonelinas (2015) in which the authors found associative priming effects for word pairs learned in the unitization condition in a lexical decision task. In line with this, we also found larger associative priming in the congruent as compared to the neutral condition in that participants made more erroneous word responses to intact than neutral compound words when the original compound words had been learned under the influence of a congruent schema versus in a schema-neutral context. Of note, as an implicit memory task as the lexical decision task does not require conscious memory judgements, the underlying memory representation is not only restricted to access by conscious memory processes (Henke, 2010).

As in the current series of experiments, schema congruency is established by means of a systematic pattern of semantic relationships, an obvious question is whether the congruency effect can be partially or completely traced back to semantic priming.

8.2 Implications

8.2.1 Are the mnemonic congruency effects the result of semantic priming?

In the current series of experiments, schema congruency is established by means of a systematic pattern of semantic relationships (see chapter 3). This circumstance could provoke the question if the associative memory benefit, which we reliably found in experiments 1 and 2, can be traced back to systematic condition differences in semantic priming of those relationships.

Within an experimental priming paradigm, a prime stimulus (usually) precedes the target and thereby, processing of the prime can influence the processing of the target stimulus, resulting in a priming effect.

Short-term priming can be distinguished from long-term priming (Wentura & Rothermund, 2014): Whereas short-term priming occurs within fractions of a second to a second, long-term priming covers longer delays between prime and target (Wentura & Rothermund, 2014). For the current work, especially semantic priming is of interest, which occurs when prime and target share semantic features. According to Wentura and Rothermund (2014), long-term semantic priming is best characterized by long-term memory changes, whereby depth of encoding (Craik & Lockhart, 1972) plays a role. Here, Wentura & Rothermund (2014) argue that a short-lived, fast-acting transition process might be accompanied by a more-enduring process, the latter resulting in changes of representational structures. From this, we can conclude that short-term semantic priming cannot account for the effects obtained from the test phase in our experiments, as short-term priming effects theoretically cannot persist for such a long time of several minutes from first to second presentation.

According to Wentura and Rothermund (2014), long-term priming includes long-term memory changes and thus might be a theoretically possible process underlying congruency effects. However, the results of experiment 1 are not in line with the idea that the associative memory benefit can be traced back to mere long-term priming effects: We found an N400 congruency effect, reflecting the additional semantic priming of the first constituent (the modifier) in the congruent as compared to the neutral condition. However, this effect was not modulated by subsequent memory success, the N400 subsequent memory effect did not differ across conditions and was therefore interpreted in that it reflects semantic integration of the constituents rather than semantic priming effects. Therefore, we assume that the congruency effects in the current work cannot be explained by priming alone but rather stem from learning under the influence of a schema.

An interesting way to directly test whether a schema is more than the sum of priming effects from its underlying semantic relationships would be to repeat experiment 1 in exchanging or modifying the definitions. Instead of showing the compound word with a definition (e.g., *A dictionary used by gardeners*$_{Congruent}$*/teachers*$_{Neutral}$ *is called...*), the compound words could be preceded with the critical word pair *gardeners teachers*, priming both constituents. Alternatively, remaining closer to the original design, the compound words should be presented with sentences like *A dictionary was used by a gardener*$_{Congruent}$*/teacher*$_{Neutral}$ hidden in small stories, to increase plausibility. If a schema is more than the sum of its semantic relationships, there should be no congruency effects in an associative memory test.

8.2.2 Schemata and unitization

Unitization has been defined as a process by means of which a holistic representation is formed out of unrelated stimuli consisting of two or more items (see Graf & Schachter, 1989). The product of this process, unitized representations, can then be treated as a single item during its processing and recognition showing up in costs (Mayes et al., 2004; 2007; Pilgrim et al., 2012) as e.g., less accessibility of the original constituents (Pilgrim et al., 2012) and also a more rigid memory representation that does not support recognition of a reversed pair (e.g., Haskins et al., 2008; Wiegand et al., 2010).

8.2.2.1 Schema-congruency can provide an entity-defining framework.

Intriguingly, it has been argued that for initial unitization to occur, an entity-defining framework is required (Mecklinger & Jäger, 2009) which in the definition-sentence paradigm is provided by the congruent definition (Bader et al., 2010). Based on the results of the current series of experiments, it appears plausible that schema-congruency can be seen as one way in which an entity-defining framework is established. This conclusion is based on the observation that experimental evidence for the formation of unitized representations has been obtained, even if the control condition (i.e., the neutral condition) does not differ in whether the word pair is processed as a compound word or not. Thus, the semantic relationship between the first constituent of the compound word (the modifier) and a noun in the definition, which leads to the embedding of both underlying concepts into a superordinate frame, seems to be crucial for whether a unitized representation is formed. Although the current series of experiments does not allow for such conclusions, it is tempting to speculate that unitization might be one mechanism accounting for how schema congruency could enable rapid neocortical learning of novel associations, as it was also argued by Hebscher et al. (2019).

8.2.2.2 The nature of the unitized memory representation

In the current thesis, we assumed that the formation of several parallel memory representations is possible (Gilboa & Moscovitch, 2021), whereby we hypothesized that schema-based learning might modulate the formation of a gist-like representation. The episodic representation might hold event-specific details which (per definition) fall through the cracks of the schema. Based on this, it might be that the episodic representation is used together with the gist-like representation and a potential viable schema representation during episodic retrieval, whilst the gist-like representation could be used alone during semantic retrieval (see Tompary et al. 2020 for a similar argument). In their model, Bastin and colleagues (2019) assume that different types of representations bring with them a preference for being used within a familiarity or a recollection process. In mapping familiarity on the gist-like unitized representation (in line with the absolute familiarity effect found in experiment 1) and recollection on a detailed episodic representation, this assumption would be in line with data from experiment 1, showing that recognition decisions regarding associations learned under the influence of a schema are supported by both, familiarity, and recollection.

The assumption that the unitized or entity representation could be gist-like does would provide an explanation for why unitized representations are rigid in that reversal of the constituents hampers recognition (Haskins et al., 2008; Wiegand et al., 2010). This is because reversal would violate the conceptual hierarchy and thereby the gist. In addition, a gist-like representation would also explain why participants underly confusion with semantically similar information, what we found in experiment 2. Semantic integration, resulting in a gist-like representation, would also be a plausible origin for the absolute familiarity signal (Bader et al., 2010; Mecklinger & Bader, 2020) which we assume to underly the N400 attenuation effect we found in experiment 1. It should be noted, however, that unitization might not always include conceptual information and thus it might not always be the case that the unitized representation constitutes a gist-like representation.

In distinguishing schema representations (that are acquired from extracting commonalities across multiple episodes, e.g., Ghosh & Gilboa, 2014; Gilboa & Marlatte, 2017; Sekeres et al., 2018) and gist-like representations that are specific for a learning instance, it can also be explained why Tibon and colleagues (2018) found no evidence for unitization in the definition condition leading to more false alarms on schema-congruent information after relearning, given that this pattern of results is not caused by methodological decisions. The schema, provided by the definition, constitutes a novel combination of an existing abstract word-formation rule and existing conceptual knowledge and might not be viable during one learning instance. Therefore, it probably cannot yet support new learning of semantically similar information. However, a gist-like representation would not be necessarily expected to be applied to congruent information during relearning. Nevertheless, it can account for higher false alarms on semantically related lures (experiment 3) in the schema-congruent condition. As we did not find higher false alarms on recombined compound words, for which the constituents are old, but the combination is new, the false alarm effect must be semantic in nature and higher false alarm rates on semantically related lure compound words cannot be explained with less precise memory representations per se.

8.2.3 Diagnostic fluency in the neurocognitive framework by Mecklinger & Bader (2020)

In their neurocognitive account of familiarity, Mecklinger and Bader (2020) consider evidence showing that familiarity is not a unitary process but rather multiply determined (e.g., Bridger et al., 2014; see Mandler, 1980). Relative familiarity is indicated by the FN400 ERP effect and is interpreted to result from the surprising difference between actually perceived and expected fluency which is fed into a mnemonic attribution process by which this discrepancy in fluency is ascribed to the learning episode. Absolute familiarity can be used in situations in which all stimuli are novel, which makes changes in absolute familiarity diagnostic for prior occurrence. In a recognition memory experiment in which novel compound words are learned with a congruent definition, the novel word pair is assigned a meaning. This meaning assignment results in semantic integration of the compound word and thereby a conceptual fluency signal. This is indicated by an N400 attenuation effect (e.g., Strózak et al., 2016; Woollams et al., 2008; Yang et al., 2019; Bruett & Leynes, 2015).

The data of the current series of experiments is in line with the neurocognitive account on familiarity in that we found the predicted N400 attenuation effect, the correlate of absolute familiarity, for the learning of novel compound words. In addition, we found that fluency can also be misused in an implicit memory task: In a lexical decision task, conceptual fluency might be diagnostic for realness of a word. The conceptual fluency of the novel compound words learned under the influence of a schema might be erroneously used in lexical decisions, leading to more false word responses on compound words in the congruent condition.

Of note, in their model, Mecklinger & Bader (2020) specify attribution processes to underly relative familiarity alone, but not absolute familiarity. At first sight, this difference seems surprising as some kind of decision processes are required for absolute familiarity to explain memory judgements. The mapping of fluency attribution to relative familiarity and the FN400 is motivated by the idea that fluency must be perceived to be unexpected (Whittlesea et al., 1998; 2001) which would require an expected baseline value to be present. As this is not the case for pairings obtaining a baseline value during the experiment, no discrepancy in fluency is to be expected.

Indirect evidence for the idea that those decision processes are at least to some extent different from the ones reflected in the FN400 ERP effect on relative familiarity comes from a study by Lloyd and colleagues (2015). In several behavioral experiments, they tested whether unitized representations could interact with fluency attribution (Jacoby & Whitehouse, 1989) in associative recognition memory tasks. Following the model by Mecklinger & Bader (2020), newly formed unitized representations are recognized by means of absolute familiarity in a recognition memory test in which all stimuli are new and conceptual fluency (absolute familiarity) is diagnostic. From this it follows that unitization should not interact with fluency attribution.

In the study by Lloyd and colleagues (2015), unitization was established by means of the definition-sentence paradigm (e.g., Bader et al., 2010; Haskins et al., 2008; Wiegand et al., 2010). In a first step, the authors showed that familiarity contributes to recognition of unitized representations in using a speeded response deadline recognition test (experiments 1a and 1b). Hereby, it is assumed that the required speed to respond should reduce the opportunity to base responses on recollection (Scheuplein et al., 2014). Thus, responses should rather rely on familiarity. As expected, memory performance in the definition condition decreased less for a speeded as compared to not speeded memory test, indicating a larger contribution of familiarity to recognition. In a next step, the authors investigated whether unitization interacts with fluency attribution. Therefore, during the recognition memory test, a priming procedure was applied in which half of the test word pairs (intact or recombined) were preceded with matching masked primes or mismatching masked primes. Matching primes resulted in the test words being presented longer whereby the first part of its presentation was masked (34ms) by a row of xxxx's. Participants were not informed about the primes and should classify each test word pair as old (intact pair) or new (recombined pair). The dependent variable was the amount of yes (old) responses. If familiarity-based remembering of the novel word pairs includes fluency attribution processes, we would expect a priming effect to occur in the definition, as compared to the sentence condition. This is because the additional conceptual fluency added by the primes should increase the discrepancy between expected and perceived fluency that is attributed to prior occurrence, resulting in familiarity. However, this was not the case, neither when unitization was manipulated between participants (experiment 2a) or within participants (experiment 2b). In their experiment 3, the authors modified the recognition memory test in that instead of recombined word pairs, new (yet unpresented) word pairs were presented. However, still no priming effects were found. The authors did not find evidence for fluency attribution occurring during recognition of unitized representations.

If the null hypothesis is statistically valid, what should be tested in future experiments, this would be in line with the neurocognitive account by Mecklinger and Bader (2020) and point towards the idea that the decision processes involved in absolute familiarity are to some extent qualitatively different to the fluency attribution processes involved in relative familiarity. Here, another aspect to consider is the phenomenological quality of memory decisions based on absolute familiarity. If no attribution to the past occurs for absolute familiarity, participants might base their judgements more on guess responses in a remember/know/guess procedure, whilst relative familiarity, e.g., for existing compound words, should lead to more know responses.

8.2.4 From novel to real compound words

An interesting question is how novel compound words, as in the present study, become real compound words, as e.g., tea mug. The reason why unitized representations are plausible format to represent compound words is their rigidity, which can protect the conceptual hierarchy necessary. The system should retain the information that a *tea mug* is a tea mug and not *mug tea*, what would violate the gist structure of the compound word.

Dual-route models of compound word processing (Isel et al., 2003, Koester et al., 2007, Koester et al., 2004, Koester et al., 2009, Sandra, 1990, Zwitserlood, 1994) assume that compound words may be represented and processed as unitary lexical units or analyzed as individual constituents via combinatorial mechanisms. As the compound words in the present study are novel, there cannot yet exist an accessible lexical entry (Libben, 2006) before the first presentation. However, the results of our experiments indicate that in situations where a novel concept emerges from the compound word, supported by relations to prior knowledge, a first holistic lexical entry, i.e., a unitized representation, may be formed only after one learning instance, enabling its familiarity-based recognition. How the newly created memory representation becomes similar to existing lexical entries is not yet known.

In a recent study, Yacoby and colleagues (2021) analyzed fMRI patterns for new pairings that are initially inconsistent with prior knowledge (e.g., green lion) and for schema consistent, existing pairings. They found that only after three repetitions, the neural patterns between inconsistent and consistent pairings could not be distinguished anymore. Following the approach by Hebscher and colleagues (2019), repetition as congruency with prior knowledge and unitization might enable a similar neocortical learning mechanism. In line with this, the holistic entries formed under the influence of a schema in our experiments might be strengthened with repeated presentations, but it is conceivable that due to their initial congruent encoding, they might require less repetitions until their neural patterns are undistinguishable from existing compound words. However, to the best of our knowledge, no study directly investigated the transition from novel compound words to real compound words. In reformulating the research question towards are more process-oriented approach, it would be interesting to test how many learning instances are required for the novel compound words to be that semantically integrated that it elicits FN400 effects of relative familiarity instead of N400 effects of absolute familiarity.

As we validated in a fourth rating study (see chapter 3.4.1), perceived semantic transparency (see Bell & Schäfer, 2016 of the compound word is closely related to schema congruency as it was operationalized within the current work. Thereby, schema congruency might enable semantic integration of both constituents into the whole word meaning, leading to a semantically transparent compound word. It is tempting to speculate that in natural language learning, compound words that are considered transparent are similarly integrated in prior knowledge structures.

8.2.5 Getting back to the neuroscientific schema approach

As we argued before, today's neuroscientific schema literature is mainly concerned with brain-systems approaches to schema-based learning (e.g., Hebscher et al., 2019; Preston & Eichenbaum, 2013; van Kesteren et al., 2012) whilst our approach was more functional in nature. However, some interesting implications can be derived from integrating the current data with the neuroscientific schema approach.

We hypothesized that in line with the neurobiological (Tse et al., 2007; 2011) and behavioral evidence (Tompary et al., 2020), schema representations might be formed in parallel to episodic representations but their expression (i.e., if they are recruited during retrieval) depends on the retrieval demands and goals (see also Tompary et al., 2020 for a similar argument; see McKenzie et al., 2013; Preston & Eichenbaum, 2013 and van Kesteren et al., 2012 for different views). We later refined this idea in that in the current series of experiments, it is not necessarily a schema representation which is formed, but rather a gist-like unitized representation, the formation of which is modulated by schema congruency. We further assume that in line with Moscovitch and Gilboa (2021), several memory representations might be formed in parallel and fed into different neurocognitive processes. This latter assumption is supported by the recognition ERP results from experiment 1, in which we found evidence for absolute familiarity and recollection to contribute to associative recognition and also by the pattern of more false alarms on semantically related lures in the congruent condition (experiment 3).

Of note, schema models differ in the role they assign to the hippocampus in schema-based learning. In the schema-linked interactions between medial prefrontal and medial temporal regions (SLIMM) model by van Kesteren and colleagues (2012), the authors assume that the mPFC and the hippocampus are reciprocally connected. If incoming information is congruent with a schema, so-called resonance is detected by the mPFC and hippocampal encoding is inhibited. As in the current work, we did not use methods allowing for conclusions on brain systems underlying our effects, we can only speculate based on our behavioral and ERP results. In assuming that similar brain regions underly encoding and retrieval of information, as it is in the model by van Kesteren and colleagues, and also assuming that the hippocampus is required for recollection to take place (Henke, 2010), the SLIMM model would have predicted to find a smaller contribution of recollection to recognition decisions when information is congruent with a schema. This is not what we found in experiment 1. We did not find statistically reliable condition differences in the left-parietal old/new effect, the putative correlate of recollection and the numerical trend was in the opposite direction. Our results are more in line with the idea that schema-based learning and traditional association learning occur in parallel (Henke, 2010). However, we found another pattern of results indicating less hippocampal involvement, which is more false alarms to semantically similar information in experiment 3. To distinguish highly similar information, hippocampal pattern separation is required (Yassa & Stark, 2011). Nevertheless, the current experiments do not permit such conclusions that should be tested in using appropriate methods in future research.

Another aspect of high importance in schema-based learning is the role of consolidation. Current neuroscientific views on rapid neocortical learning, driven by congruency with prior knowledge, propose at least two ways by which a neocortical engram can be formed and expressed, i.e., fast integration during encoding (Lesburguères et al., 2011; van Kesteren et al., 2012) and rapid consolidation (Tse et al., 2007). Whilst we found evidence for the contribution of different encoding processes underlying later successful recognition of novel schema-congruent compound words in experiment 1 of the current work, we cannot draw any conclusions about the contribution of consolidation to schema-based learning of associations based on our experiments.

Following the standard consolidation theory, it is believed that reorganizational processes render initially hippocampus-dependent memory engrams hippocampus-independent, leaving behind neocortical representations (Dudai, 2012). Sleep is considered to be beneficial for systems consolidation to occur (Dudai, 2012). Regarding fast mapping, another form of learning involving prior knowledge (see Hebscher et al., 2019), Himmer et al. (2017) found that integration might already occur during encoding, as associations acquired via fast mapping benefitted less from consolidation during sleep, than associations that were explicitly encoded, which is in line with fast mapping initiating the rapid formation of neocortical representations.

This is of special interest for our research on compound word learning, as it remains unclear if the schema exerts its influence by means of mere fast integration during encoding, probably reflected in our ERP effects, or rapid consolidation. In a follow-up experiment, we would like to test if rapid integration takes place for compound words initially learned in the congruent condition, resulting in memory performance benefitting less from sleep-enabled consolidation than in the neutral condition.

8.2.6 Why is incongruency not better than congruency for compound word learning?

In the literature on schema-based learning, there is evidence for both a congruency effect (e.g., Bein et al., 2014; 2015), i.e., better memory for congruent information but also evidence for an incongruency effect (e.g., Greve et al., 2019). Greve and colleagues 2019 show that both effects can occur in the same experiments, but different processes underly both phenomena. There are several reasons why there are probably no incongruency effects (i.e., a memory advantage for the neutral as compared to the congruent condition) in the current experiments. First, it is questionable whether the neutral condition is comparable to an incongruent condition as in Greve and colleagues (2019). In natural language processing, opaque compound words exist (e.g., Libben, 2006), for which the contribution of at least one constituent to the meaning of the whole compound word remains unclear, e.g., elderberry. Thus, especially for compound words, the neutral condition is probably better interpreted as an information gap than real incongruency. Second, the incongruency effect is usually attributed to hippocampal processing (van Kesteren et al., 2012). The capacity for episodic learning might be limited (Harkotte et al., 2022). However, learning that many new pairings introduces a high information load. In a recent rodent study, Harkotte and colleagues (2022), manipulated information load (high, low) and the type of memory tested (schematic, episodic) in a single experiment. They found that high versus low information load was beneficial for schematic memories whilst episodic memories were better for low as compared to high information load. From this it is conceivable that high information load in the current experiments was beneficial for relying on schema-based learning. However, this hypothesis should be tested in follow-up experiments.

8.3 Limitations and Future directions

In the current experiment, we found first evidence for an extended N400 effect during encoding and also later recognition of novel compound words. We discussed that in contrast to previous compound word learning experiments, compound words in the current experiments were presented as a whole, without a blank separating the constituents. In a previous eye tracking study, the presence or absence of such a blank space between constituents has been found to influence speed of processing of the compound word and also how well a holistic entry could be formed (Inhoff et al., 2000). However, further research is required to shed light on the validity of the extended N400 effect, which is of particular interest for the interpretation of more absolute familiarity underlying recognition in the congruent as compared to the neutral condition. The latter aspect should be verified in follow-up experiments using behavioral measures to disentangle familiarity and recollection either in concert with an ERP approach or within mere behavioral experiments. Here, either the remember/know procedure (Gardiner, 1988; Tulving, 1985) could be applied to obtain familiarity estimates on a single-trial level or ROC curves (Yonelinas et al., 1999) to obtain per-condition-estimates of the contribution of familiarity to recognition. In addition, it would be interesting to test which neurocognitive processes underly false alarms on semantically similar information (see experiment 3). More precisely, it is tempting to speculate that false alarms on semantically similar information in the schema congruent condition should be accompanied by an absolute familiarity ERP effect similar to the effect on hits. However, it is also possible that phantom recollection (e.g., Dennis et al., 2012) underlies false the false alarm effect on semantically similar information.

In addition, it would also be interesting to shed light on the nature of the extended N400 effect during initial processing of novel compound words. In another experiment, the presence or absence of a blank space between constituents could be systematically manipulated. For this experiment, it would be sufficient to repeat the initial learning phase without testing later memory to check if there are systematic differences in the latency and duration of the N400 component. Another possibility is that meaning integration of the novel compound word occurs only when the head is processed. If bottom-up semantic integration of the constituents is beneficial for memory formation and underlies the N400-SME in both conditions (see Discussion experiment 1), then we should find a similar N400-SME only on the second, i.e., the head constituent in an auditory version of the experiment, in which constituent processing is sequential in nature.

Another aspect awaiting testing is the role of consolidation in schema-based learning of novel compound words (see above)

9 Conclusion

In the current work, we were interested in a functional perspective on schema-based learning of novel associations. In combining the idea that schema-congruency initiates a specific learning mechanism relying less on hippocampal contribution (e.g., Hebscher et al., 2019; van Kesteren et al., 2012) and that associations can be learned with less hippocampal contribution when unitized representations are formed (e.g., Bader et al., 2014; Haskins et al., 2008; see Henke, 2010), we hypothesized that schema-congruency might support the formation of unitized representations. Hereby, the existence of a unitized representation resulting from semantic integration of the compound word might result in conceptual fluency, which is diagnostic in a recognition memory test (Mecklinger & Bader, 2020).

In the current series of experiments, we found first support for the idea that schema congruency enables the formation of unitized representations (experiment 1 and 3) and leads to absolute familiarity contributing to recognition memory judgements (experiment 1). Further, we found evidence in that whether schema-congruency is beneficial for associative memory performance depends on the nature of the memory test. If associative information is tested by recombining item information, violating the novel concept learned, schema-congruence is beneficial for memory performance (experiment 1 and 3). However, if semantically similar associative information is tested, fulfilling the novel concept, schema congruency is detrimental for memory performance, leading to more false alarms in semantically similar information (experiment 3). The results were interpreted in that the unitized representation might be gist-like.

In referring back to the research questions formulated in chapter 2.4, we interpreted our results in that the neurocognitive mechanisms underlying schema-based learning include semantic priming, establishing schema congruency, semantic integration of the constituents and the formation of a conceptual (unitized) representation whereby evidence for semantic integration being beneficial for memory was independent from condition.

Second, the memory representations formed and the processes underlying later recognition include unitization, and absolute familiarity contributing to associative recognition. The representations formed under the influence of a schema include a gist-like, unitized representation next to episodic associations as they are formed in traditional associative learning. Lastly, those representations cannot only be retrieved in an explicit memory test, but also influence performance in an implicit memory test, supporting the idea that consciousness does not necessarily play a role for how memory representations are stored (Henke, 2010; Reder et al., 2009).

Reference list

Ahmad, F. N., & Hockley, W. E. (2014). The role of familiarity in associative recognition of unitized compound word pairs. *Quarterly Journal of Experimental Psychology, 67*(12), 2301–2324. https://doi.org/10.1080/17470218.2014.923007

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93*(2), 203–231.

Alonso, A., van der Meij, J., Tse, D., & Genzel, L. (2020). Naïve to expert: Considering the role of previous knowledge in memory. *Brain and Neuroscience Advances, 4*, 1-17. https://doi.org/10.1177/2398212820948686

Anderson, M. C. (2020). Episodic Memory: Organizing and remembering. In In A. Baddeley; M.W. Eysenck; M.C. Anderson *Memory* (3rd Edition). (S. 163–205). Routledge.

Atienza, M., Crespo-Garcia, M., & Cantero, J. L. (2011). Semantic congruence enhances memory of episodic associations: Role of theta oscillations. *Journal of Cognitive Neuroscience, 23*(1), 75–90. https://doi.org/10.1162/jocn.2009.21358

Baddeley, A. (2020). What is memory? In In A. Baddeley; M.W. Eysenck; M.C. Anderson *Memory* (3rd Edition). (S. 3–21). Routledge.

Bader, R. (2014). *Using unitization as encoding strategy in associative recognition memory: Behavioral, fMRI, and ERP evidence.* Saarbrücken, 2014.

Bader, R., & Mecklinger, A. (2017). Separating event-related potential effects for conceptual fluency and episodic familiarity. *Journal of Cognitive Neuroscience*, *29*(8), 1402–1414. https://doi.org/10.1162/jocn_a_01131

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

Bader, R., Mecklinger, A., & Meyer, P. (2020). Usefulness of familiarity signals during recognition depends on test format: Neurocognitive evidence for a core assumption of the CLS framework.

Neuropsychologia, 148, 107659. https://doi.org/10.1016/j.neuropsychologia.2020.107659

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

Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Di Russo, F. (2016). Disentangling metaphor from context: An ERP study. *Frontiers in Psychology, 7*, Article 559. https://doi.org/10.3389/fpsyg.2016.00559

Barry, D. N., & Maguire, E. A. (2019). Remote memory and the hippocampus: A constructive critique. *Trends in Cognitive Sciences, 23*(2), 128–142. https://doi.org/10.1016/j.tics.2018.11.005

Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge University Press.

Bastin, C., Besson, G., Simon, J., Delhaye, E., Geurten, M., Willems, S., & Salmon, E. (2019). An integrative memory model of recollection and familiarity to understand memory deficits. *Behavioral and Brain Sciences, 42*, e281, 1-60. https://doi.org/10.1017/S0140525X19000621

Bein, O., Livneh, N., Reggev, N., Gilead, M., Goshen-Gottstein, Y., & Maril, A. (2015). Delineating the effect of semantic congruency on episodic memory: The role of integration and relatedness. *PLOS ONE, 10*(2), Article e0115624. https://doi.org/10.1371/journal.pone.0115624

Bein, O., Reggev, N., & Maril, A. (2014). Prior knowledge influences on hippocampus and medial prefrontal cortex interactions in subsequent memory. *Neuropsychologia, 64*, 320–330. https://doi.org/10.1016/j.neuropsychologia.2014.09.046

Bell, M. J., & Schäfer, M. (2016). Modelling semantic transparency. *Morphology, 26*(2), 157–199. https://doi.org/10.1007/s11525-016-9286-3

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). Academic Press.

Boutonnet, B., McClain, R., & Thierry, G. (2014). Compound words prompt arbitrary semantic associations in conceptual memory. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00222

Bridger, E. K., Bader, R., Kriukova, O., Unger, K., & Mecklinger, A. (2012). The FN400 is functionally distinct from the N400. *NeuroImage*, *63*(3), 1334–1342. https://doi.org/10.1016/j.neuroimage.2012.07.047

Bridger, E. K., Bader, R., & Mecklinger, A. (2014). More ways than one: ERPs reveal multiple familiarity signals in the word frequency mirror effect. *Neuropsychologia*, *57*, 179–190. https://doi.org/10.1016/j.neuropsychologia.2014.03.007

Brothers, T., Wlotko, E. W., Warnke, L., & Kuperberg, G. R. (2020). Going the extra mile: Effects of discourse context on two late positivities during language comprehension. *Neurobiology of Language*, *1*(1), 135–160. https://doi.org/10.1162/nol_a_00006

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about Semantic Illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143. https://doi.org/10.1016/j.brainres.2012.01.055

Bruett, H., & Leynes, P. A. (2015). Event-related potentials indicate that fluency can be interpreted as familiarity. *Neuropsychologia*, *78*, 41–50. https://doi.org/10.1016/j.neuropsychologia.2015.09.035

Butterworth, B. (1983). Lexical Representation. In *Language production*, vol. 2 (S. 257–294). Academic Press Inc. (London) Ltd.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, *28*, 297–332.

Champely S (2020). pwr: Basic Functions for Power Analysis. R package version 1.3-0, <https://CRAN.R-project.org/package=pwr>.

Cockcroft, J. P., Berens, S. C., Gaskell, M. G., & Horner, A. J. (2022). Schematic information influences memory and generalisation behaviour for schema-relevant and -irrelevant information. *Cognition, 227,* 105203. https://doi.org/10.1016/j.cognition.2022.105203

Cohen, N., Pell, L., Edelson, M. G., Ben-Yakov, A., Pine, A., & Dudai, Y. (2015). Peri-encoding predictors of memory encoding and consolidation. *Neuroscience & Biobehavioral Reviews, 50*, 128–142. https://doi.org/10.1016/j.neubiorev.2014.11.002

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407–428. https://doi.org/10.1037/11571-006

Coulson, S., & Kutas, M. (2001). Getting it: Human event-related brain response to jokes in good and poor comprehenders. *Neuroscience Letters, 316*(2), 71–74. https://doi.org/10.1016/S0304-3940(01)02387-4

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology, 16,* 693-700. https://doi.org/10.1016/j.conb.2006.10.012

Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences, 100*(4), 2157–2162. https://doi.org/10.1073/pnas.0337195100

De Brigard, F. (2014). The Nature of Memory Traces. *Philosophy Compass*, *9*(6), 402–414. https://doi.org/10.1111/phc3.12133

de Chastelaine, M., Mattson, J. T., Wang, T. H., Donley, B. E., & Rugg, M. D. (2016). The neural correlates of recollection and retrieval monitoring: Relationships with age and recollection performance. *NeuroImage*, *138,* 164–175. https://doi.org/10.1016/j.neuroimage.2016.04.071

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117–1121. https://doi.org/10.1038/nn1504

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dennis, N. A., Bowman, C. R., & Vandekar, S. N. (2012). True and phantom recollection: An fMRI investigation of similar and distinct neural correlates and connectivity. *NeuroImage, 59*(3), 2982–2993. https://doi.org/10.1016/j.neuroimage.2011.09.079

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

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends in Cognitive Sciences, 11*(9), 379–386. https://doi.org/10.1016/j.tics.2007.08.001

Dien, J., & Santuzzi, A. M. (2005). Application of repeated measures ANOVA to high-density ERP datasets: A review and tutorial. In T. C. Handy (Hrsg.), *Event-related potentials: A methods handbook* (S. 1–73). MIT Press.

Dudai, Y. (2012). The Restless Engram: Consolidations Never End. *Annual Review of Neuroscience, 35*(1), 227–247. https://doi.org/10.1146/annurev-neuro-062111-150500

Düzel, E., Vargha-Khadem, F., Heinze, H. J., & Mishkin, M. (2001). Brain activity evidence for recognition without recollection after early hippocampal damage. *Proceedings of the National Academy of Sciences, 98*(14), 8101–8106. https://doi.org/10.1073/pnas.131205798

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30*(1), 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Fabiani, M., & Donchin, E. (1995). Encoding processes and memory organization: A model of the von Restorff Effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 224–240.

Fabiani, M., Karis, D., & Donchin, E. (1986). P300 and recall in an incidental memory paradigm. *Psychophysiology, 23*(3), 298–308.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*(4), 491–505. https://doi.org/10.1111/j.1469-8986.2007.00531.x

Field, A. (2009). *Discovering Statistics Using IBM SPSS Statistics*. SAGE Publications.

Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: a unified JavaScript framework for online studies. In International Conference on Computational Social Science (Cologne).

Fiske, S. T., & Linville, P. W. (1980). What does the schema concept buy us? *Personality and Social Psychology Bulletin, 6*(4), 543–557. https://doi.org/10.1177/014616728064006

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience, 6*(2), 119–130. https://doi.org/10.1038/nrn1607

Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research, 31*(1), 45–63.

Friedman, D., & Johnson, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique, 51,* 6–28.

Gagné, C. L. (2002). Lexical and relational influences on the processing of novel compounds. *Brain and Language, 81*(1–3), 723–735. https://doi.org/10.1006/brln.2001.2559

Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(1), 71–87.

Gagné, C. L., & Spalding, T. L. (2004). Effect of discourse context and modifier relation frequency on conceptual combination. *Journal of Memory and Language, 50*(4), 444–455. https://doi.org/10.1016/j.jml.2004.01.003

Gagné, C. L., & Spalding, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language, 60*(1), 20–35. https://doi.org/10.1016/j.jml.2008.07.003

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition, 16*(4), 309–313. https://doi.org/10.3758/BF03197041

Gerrig, R. J., & Bortfeld, H. (1999). Sense Creation in and out of Discourse Contexts. *Journal of Memory and Language, 41*(4), 457–468. https://doi.org/10.1006/jmla.1999.2656

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

Gibney, K. D., Kypriotakis, G., Cinciripini, P. M., Robinson, J. D., Minnix, J. A., & Versace, F. (2020). Estimating statistical power for event-related potential studies using the late positive potential. *Psychophysiology, 57,* e13482. https://doi. org/10.1111/psyp.13482

Gilboa, A., & Marlatte, H. (2017). Neurobiology of Schemas and Schema-Mediated Memory. *Trends in Cognitive Sciences, 21*(8), 618–631. https://doi.org/10.1016/j.tics.2017.04.013

Gilboa, A., & Moscovitch, M. (2021). No consolidation without representation: Correspondence between neural and psychological representations in recent and remote memory. *Neuron, 109*(14), 2239–2255. https://doi.org/10.1016/j.neuron.2021.04.025

Giner-Sorolla, R., Montoya, A. K., Aberson, C., Carpenter, T., Lewis, N. A., Bostyn, D. H., Conrique, B., Ng, B. W., Reifman, A., Schoemann, A. M., & Soderberg, C. K. (2023). Power to detect what? Considerations for planning and evaluating sample size [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/rv3kw

Gonsalves, B., & Paller, K. A. (2000). Neural Events that underlie remembering something that never happened. *Nature Neuroscience, 3*(12), 1316–1321.

Graesser, A. C., Gordon, S. E., & Sawyer, J. D. (1979). Recognition memory for typical and atypical actions in scripted activities: Tests of a script pointer + tag hypothesis. *Journal of Verbal Learning and Verbal Behavior*, *18*(3), 319–332. https://doi.org/10.1016/S0022-5371(79)90182-8

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(3), 501–518.

Graf, P., & Schacter, D. L. (1989). Unitization and grouping mediate dissociations in memory for new associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 930–940. https://doi.org/10.1037/0278-7393.15.5.930

Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. *Journal of Experimental Psychology: General*, *148*(2), 325–341. https://doi.org/10.1037/xge0000498

Greve, A., van Rossum, M. C. W., & Donaldson, D. I. (2007). Investigating the functional interaction between semantic and episodic memory: Convergent behavioral and electrophysiological evidence for the role of familiarity. *NeuroImage*, *34*(2), 801–814. https://doi.org/10.1016/j.neuroimage.2006.07.043

Grosjean, P., Ibanez, F. (2018). pastecs: Package for Analysis of Space-Time Ecological Series. R package version 1.3.21, <https://CRAN.R-project.org/package=pastecs>.

Hagoort, P. (2003). Interplay between syntax and Semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, *15*(6), 883–899. https://doi.org/10.1162/089892903322370807

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (sps) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8*(4), 439–483. https://doi.org/10.1080/01690969308407585

Hannula, D. E., & Ranganath, C. (2008). Medial temporal lobe activity predicts successful relational memory binding. *The Journal of Neuroscience, 28*(1), 116–124. https://doi.org/10.1523/JNEUROSCI.3086-07.2008

Harkotte, M., Contreras, M. P., Inostroza, M., & Born, J. (2022). Effects of information load on schema and episodic memory formation. *Frontiers in Behavioral Neuroscience, 16*, 923713. https://doi.org/10.3389/fnbeh.2022.923713

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

Häuser, K. I., Demberg, V., & Kray, J. (2019). Effects of aging and dual-task demands on the comprehension of less expected sentence continuations: Evidence from pupillometry. *Frontiers in Psychology, 10*, Article 709. https://doi.org/10.3389/fpsyg.2019.00709

Head, B. Y. H., & Holmes, G. (1911). Sensory disturbances from cerebral lesions. *Brain, 34*(2), 103–254.

Hebscher, M., Wing, E., Ryan, J., & Gilboa, A. (2019). Rapid cortical plasticity supports long-term memory formation. *Trends in Cognitive Sciences, 23*(12), 989–1002. https://doi.org/10.1016/j.tics.2019.09.009

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). DlexDB – eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau, 62*(1), 10–20. https://doi.org/10.1026/0033-3042/a000029

Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience, 11*(7), 523–532. https://doi.org/10.1038/nrn2850

Himmer, L., Müller, E., Gais, S., & Schönauer, M. (2017). Sleep-mediated memory consolidation depends on the level of integration at encoding.

Neurobiology of Learning and Memory, 137, 101–106. https://doi.org/10.1016/j.nlm.2016.11.019

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology, 30*(1), 47–61. https://doi.org/10.1111/j.1469-8986.1993.tb03204.x

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70.

Höltje, G., Lubahn, B., & Mecklinger, A. (2019). The congruent, the incongruent, and the unexpected: Event-related potentials unveil the processes involved in schematic encoding. *Neuropsychologia, 131,* 285–293. https://doi.org/10.1016/j.neuropsychologia.2019.05.013

Höltje, G., & Mecklinger, A. (2022). Benefits and costs of predictive processing: How sentential constraint and word expectedness affect memory formation. *Brain Research, 1788*, 147942. https://doi.org/10.1016/j.brainres.2022.147942

Hubbard, R. J., Rommers, J., Jacobs, C. L., & Federmeier, K. D. (2019). Downstream behavioral and electrophysiological consequences of word prediction on recognition memory. *Frontiers in Human Neuroscience, 13*, 291. https://doi.org/10.3389/fnhum.2019.00291

Huffer, V., Bader, R., & Mecklinger, A. (2022). Can the elderly take the action? – The influence of unitization induced by action relationships on the associative memory deficit. *Neurobiology of Learning and Memory*, 107655. https://doi.org/10.1016/j.nlm.2022.107655

Ionita, C., Talmi, D., & Taylor, J. (2019). Priming recognition memory test cues: No evidence for an attributional basis of recollection. *Behavioral and Brain Sciences, 42,* E289. https://doi.org/doi:10.1017/S0140525X19001948

Iran-Nejad, A. (1980). The schema: A structural or a functional pattern. (Technical Report No. 159). Urbana: University of Illinois, Center for the Study of Reading. (ERIC Document Reproduction Service No. ED 182 7.

Iran-Nejad, A., & Winsler, A. (2000). Bartlett's Schema Theory and Modern Accounts of Learning and Remembering. *The Journal of Mind and Behavior, 21*(1/2), 5–35.

Isel, F., Gunter, T. C., & Friederici, A. D. (2003). Prosody-assisted head-driven access to spoken German compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(2), 277–288. https://doi.org/10.1037/0278-7393.29.2.277

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*(3), 306–340.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In *Essays in honor of Endel Tulving* (S. 391–422). Lawrence Erlbaum.

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General, 118*(2), 126–135.

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

Jasper, H. (1958). Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalography and Clinical Neurophysiology, 10*(2), 370–375. https://doi.org/10.1016/0013-4694(58)90053-1

Josselyn, S. A., & Tonegawa, S. (2020). Memory engrams: Recalling the past and imagining the future. *Science, 367*(6473), eaaw4325. https://doi.org/10.1126/science.aaw4325

Kamp, S.-M. (2020). Neurocognitive mechanisms of guided item and associative encoding in young and older adults. *Brain and Cognition, 145*, 105626. https://doi.org/10.1016/j.bandc.2020.105626

Kamp, S.-M., Bader, R., & Mecklinger, A. (2016). The Effect of Unitizing Word Pairs on Recollection Versus Familiarity-Based Retrieval—Further Evidence From ERPs. *Advances in Cognitive Psychology, 12*(4), 168–177. https://doi.org/10.5709/acp-0196-2

Kamp, S.-M., Bader, R., & Mecklinger, A. (2017). ERP Subsequent Memory Effects Differ between Inter-Item and Unitization Encoding Tasks. *Frontiers in Human Neuroscience, 11,* Article 30. https://doi.org/10.3389/fnhum.2017.00030

Kamp, S.-M., Potts, G. F., & Donchin, E. (2015). On the roles of distinctiveness and semantic expectancies in episodic encoding of emotional words: Emotional word encoding. *Psychophysiology*, *52*(12), 1599–1609. https://doi.org/10.1111/psyp.12537

Karis, D., Fabiani, M., & Donchin, E. (1984). "P300" and memory: Individual differences in the von Restorff effect. *Cognitive Psychology*, *16*(2), 177–216. https://doi.org/10.1016/0010-0285(84)90007-0

Keltner, D., & Haidt, J. (2003). Approaching awe, a moral, spiritual, and aesthetic emotion. *Cognition and Emotion*, *17*(2), 297–314. https://doi.org/10.1080/02699930302297

Koester, D., Gunter, T. C., & Wagner, S. (2007). The morphosyntactic decomposition and semantic composition of German compound words investigated by ERPs. *Brain and Language*, *102*(1), 64–79. https://doi.org/10.1016/j.bandl.2006.09.003

Koester, D., Gunter, T. C., Wagner, S., & Friederici, A. D. (2004). Morphosyntax, Prosody, and Linking Elements: The Auditory Processing of German Nominal Compounds. *Journal of Cognitive Neuroscience*, *16*(9), 1647–1668. https://doi.org/10.1162/0898929042568541

Koester, D., Holle, H., & Gunter, T. C. (2009). Electrophysiological evidence for incremental lexical-semantic integration in auditory compound comprehension. *Neuropsychologia*, *47*, 1854–1864. https://doi.org/10.1016/j.neuropsychologia.2009.02.027

Kriukova, O., Bridger, E., & Mecklinger, A. (2013). Semantic relations differentially impact associative recognition memory: Electrophysiological evidence. *Brain and Cognition*, *83*(1), 93–103. https://doi.org/10.1016/j.bandc.2013.07.006

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A Tale of Two Positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience, 32*(1), 12–35. https://doi.org/10.1162/jocn_a_01465

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4,* Article 863.

Lashley, K. S. (1950). In search of the engram. In In J. F. Danielli, R. Brown, Eds. *Society for Experimental Biology, Physiological mechanisms in animal behavior.* (Society's Symposium IV.) (S. 454–482). Academic Press.

Lechner, H. A., Squire, L. R., & Byrne, J. H. (1999). 100 years of consolidation—Remembering Müller and Pilzecker. *Learning & Memory, 6*(2), 77–87. https://doi.org/10.1101/lm.6.2.77

Leiner, D. J. (2019). SoSci Survey (Version 3.1.06) [Computer software]. Available at https://www.soscisurvey.de.

Lesburguères, E., Gobbo, O. L., Alaux-Cantin, S., Hambucken, A., Trifilieff, P., & Bontempi, B. (2011). Early tagging of cortical networks is required for the formation of enduring associative memory. *Science, 331*(6019), 924–928. https://doi.org/10.1126/science.1196164

Leynes, P. A., & Bink, M. L. (2002). Did I do that? An ERP study of memory for performed and planned actions. *International Journal of Psychophysiology, 45*(3), 197–210. https://doi.org/10.1016/S0167-8760(02)00012-0

Libben, G. (2006). Why study compound processing? An overview of the issues. In *The Representation and Processing of Compound Words.* (S. 1–22). Oxford Linguistics.

Lloyd, M. E., Hartman, A., Ngo, C. T., Ruser, N., Westerman, D. L., & Miller, J. K. (2015). Not enough familiarity for fluency: Definitional encoding increases familiarity but does not lead to fluency attribution in associative recognition. *Memory & Cognition, 43*(1), 39–48. https://doi.org/10.3758/s13421-014-0449-3

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8*, Article 213. https://doi.org/10.3389/fnhum.2014.00213

Luck, S. J. (2014). An introduction to the event-related potential technique (Second edition). The MIT Press.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology, 54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Maguire, E. A. (2022). Does memory research have a realistic future? *Trends in Cognitive Sciences*, *26*(12), 1043–1046. https://doi.org/10.1016/j.tics.2022.07.006

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*(3), 252–271. https://doi.org/10.1037/0033-295X.87.3.252

Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and Meaning in the English Mental Lexicon. *Psychological Review, 101*(1), 3–33.

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

Mayes, A. R., Holdstock, J. S., Isaac, C. L., Montaldi, D., Grigor, J., Gummer, A., Cariga, P., Downes, J. J., Tsivilis, D., Gaffan, D., Gong, Q., & Norman, K. A. (2004). Associative recognition in a patient with selective hippocampal lesions and relatively normal item recognition. *Hippocampus, 14*(6), 763–784. https://doi.org/10.1002/hipo.10211

McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and clinical Neurophysiology, 62*, 203–208.

McClelland, J. L. (2013). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal of Experimental Psychology: General, 142*(4), 1190–1210. https://doi.org/10.1037/a0033812

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex:Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419–457.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In In G. H. Bower (Ed.), *The psychology of learning and motivation* (Bd. 24, S. 109–165). Academic Press.

McKenzie, S., Frank, A. J., Kinsky, N. R., Porter, B., Rivière, P. D., & Eichenbaum, H. (2014). Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron, 83*(1), 202–215. https://doi.org/10.1016/j.neuron.2014.05.019

McVee, M. B., Dunsmore, K., & Gavelek, J. R. (2005). Schema theory revisited. *Review of Educational Research, 75*(4), 531–566. https://doi.org/10.3102/00346543075004531

Mecklinger, A., & Bader, R. (2020). From fluency to recognition decisions: A broader view of familiarity-based remembering. *Neuropsychologia, 146*, 107527. https://doi.org/10.1016/j.neuropsychologia.2020.107527

Mecklinger, A., & Jäger, T. (2009). Episodic memory storage and retrieval: Insights from electrophysiological measures. In *Neuroimaging and Psychological Theories of Human Memory* (S. 357–382). Oxford University Press.

Mecklinger, A., & Müller, N. (1996). Dissociations in the processing of "What" and "Where" information in working memory: An event-related potential analysis. *Journal of Cognitive Neuroscience*, 8(5), 453–473. https://doi.org/10.1162/jocn.1996.8.5.453

Mecklinger, A., von Cramon, D. Y., & Matthes-von Cramon, G. (1998). Event-related potential evidence for a specific recognition memory deficit in adult survivors of cerebral hypoxia. *Brain, 121*(10), 1919–1935. https://doi.org/10.1093/brain/121.10.1919

Meßmer, J. A., Bader, R., & Mecklinger, A. (2021). The more you know: Schema-congruency supports associative encoding of novel compound words. Evidence from event-related potentials. *Brain and Cognition, 155,* 105813. https://doi.org/10.1016/j.bandc.2021.105813

Meßmer, J. A., Bader, R., & Mecklinger, A. The more you now: Schema-congruency supports the formation of unitized representations. Evidence from event-related potentials. [Manuscript submitted for publication]. Experimental Neuropsychology Unit, Saarland University

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227–234. https://doi.org/10.1037/h0031564

Middleton, E. L., Rawson, K. A., & Wisniewski, E. J. (2011). How do we process novel conceptual combinations in context? *Quarterly Journal of Experimental Psychology, 64*(4), 807–822. https://doi.org/10.1080/17470218.2010.520414

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*(5), 519–533. https://doi.org/10.1016/S0022-5371(77)80016-9

Müller, G. E., & Pilzecker, A. (1900). Experimentelle Beiträge zur Lehre vom Gedächtnis. *Psychol. Ergänzungsband 1:* 1–300.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1170–1187. https://doi.org/10.1037/0278-7393.26.5.1170

Neisser, U. (1967). *Cognitive Psychology*. Prentice-Hall, Inc.

Neville, H. J., Kutas, M., Chesney, G., & Schmidt, A. L. (1986). Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. *Journal of Memory and Language, 25*(1), 75–92.

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review, 110*(4), 611–646. https://doi.org/10.1037/0033-295X.110.4.611

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*(6), 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z

Osterhout, L., McKinnon, R., Bersick, M., & Corey, V. (1996). On the language specificity of the brain response to syntactic anomalies: Is the syntactic positive shift a member of the P300 family? *Journal of Cognitive Neuroscience*, *8*(6), 507–526. https://doi.org/10.1162/jocn.1996.8.6.507

Otten, L. J., & Donchin, E. (2000). Relationship between P300 amplitude and subsequent recall for distinctive events: Dependence on type of distinctiveness attribute. *Psychophysiology, 37*, 644–661.

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

Paller, K. A., Kutas, M., & Mayes, R. (1987). Neural correlates of encoding in an incidental learning paradigm. *Electroencephalography and Clinical Neurophysiology, 67*, 360–371.

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

Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Sciences*, *6*(2), 93–102. https://doi.org/10.1016/S1364-6613(00)01845-3

Parks, C. M., & Yonelinas, A. P. (2015). The importance of unitization for familiarity-based learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 881–903. https://doi.org/10.1037/xlm0000068

Parks, T., & McClellan, J. (1972). Chebyshev Approximation for Nonrecursive Digital Filters with Linear Phase. *IEEE Transactions on Circuit Theory*, *19*(2), 189–194. https://doi.org/10.1109/TCT.1972.1083419

Payne, B. R., Ng, S., Shantz, K., & Federmeier, K. D. (2020). Event-related brain potentials in multilingual language processing: The N's and P's. In *Psychology of Learning and Motivation* (Bd. 72, S. 75–118). Elsevier. https://doi.org/10.1016/bs.plm.2020.03.003

Piaget, J. (1926). *The language and thought of the child*. Kegan Paul, Trench, Trubner & Co. LTD.

Pichert, J. W., & Anderson, R. C. (1977). Taking different perspectives on a ttory. *Journal of Educational Psychology*, *69*(4), 309–315.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*(2), 127–152. https://doi.org/10.1111/1469-8986.3720127

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

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, *23*(17), R764–R773. https://doi.org/10.1016/j.cub.2013.05.041

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

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ranganath, C. (2022). What is episodic memory and how do we use it? *Trends in Cognitive Sciences 26*(12),1056-1061. https://doi.org/10.1016/j.tics.2022.09.023

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience, 13*(10), 713–726. https://doi.org/10.1038/nrn3338

Ratcliff, R., Gronlund, S. D., & Sheu, C.-F. (1992). Testing global memory models using ROC curves. *Psychological Review, 99*(3), 518-535.

Reder, L. M., Park, H., & Kieffaber, P. D. (2009). Memory systems do not divide on consciousness: Reinterpreting memory in terms of activation and binding. *Psychological Bulletin, 135*(1), 23–49. https://doi.org/10.1037/a0013974

Regel, S., Gunter, T. C., & Friederici, A. D. (2011). Isn’t it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience, 23*(2), 277–293.

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

Rhodes, S. M., & Donaldson, D. I. (2008). Electrophysiological evidence for the effect of interactive imagery on episodic memory: Encouraging familiarity for non-unitized stimuli during associative recognition.

NeuroImage, 39(2), 873–884. https://doi.org/10.1016/j.neuroimage.2007.08.041

Roediger, H. L., & McDermott, K. B. (1995). Creating False Memories: Remembering Words Not Presented in Lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803–814.

RStudio Team. (2019). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

RStudio Team (2022). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Rudy, J. W., & Sutherland, R. J. (2008). Is it systems or cellular consolidation? Time will tell. An alternative interpretation of the Morris group's recent science paper. *Neurobiology of Learning and Memory, 89*(4), 366–369. https://doi.org/10.1016/j.nlm.2007.07.017

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

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In *Theoretical issues in reading comprehension. Perspectives from cognitive psychology, linguistics, artificial intelligence, and education*. Lawrence Erlbaum.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In *Schooling and the acquisition of knowledge* (S. 99–135). Lawrence Erlbaum.

Sandra, D. (1990). On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *The Quarterly Journal of Experimental Psychology Section A, 42*(3), 529–567. https://doi.org/10.1080/14640749008401236

Sanquist, T. F., Rohrbauch, J. W., Syndulko, K., & Lindsley, D. B. (1980). Electrocortical signs of levels of processing: Perceptual anaysis and recognition memory. *Psychophysiology, 17*(6), 568–576.

Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: Single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and Language, 137*, 29–39. https://doi.org/10.1016/j.bandl.2014.07.010

Schacter, D. L., Dobbins, I. G., & Schnyer, D. M. (2004). Specificity of priming: A cognitive neuroscience perspective. *Nature Reviews Neuroscience, 5*(11), 853–862. https://doi.org/10.1038/nrn1534

Schank, R. C. (1982). *Dynamic memory: A theory of reminding and learning in computers and people.* New York: Cambridge University Press.

Scheuplein, A.-L., Bridger, E. K., & Mecklinger, A. (2014). Is faster better? Effects of response deadline on ERP correlates of recognition memory in younger and older adults. *Brain Research, 1582*, 139–153. https://doi.org/10.1016/j.brainres.2014.07.025

Schlichting, M. L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, 1, 1–8. https://doi.org/10.1016/j.cobeha.2014.07.005

Schulman, A. I. (1974). Memory for words recently classified. *Memory & Cognition, 2*(1A), 47–52. https://doi.org/10.3758/BF03197491

Schulte im Walde, S., & Borgwaldt, S. R. (2015). Association norms for German noun compounds and their constituents. *Behavior Research Methods, 47*(4), 1199–1221. https://doi.org/10.3758/s13428-014-0539-y

Sekeres, M. J., Winocur, G., & Moscovitch, M. (2018). The hippocampus and related neocortical structures in memory transformation. *Neuroscience Letters, 680*, 39–53. https://doi.org/10.1016/j.neulet.2018.05.006

Semon, R. (1904). *Die Mneme als erhaltendes Prinzip im Wechsel des organischen Geschehens.* W. Engelmann, Ed.

Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M. (2022). afex: Analysis of Factorial Experiments. R package version 1.2-0, <https://CRAN.R-project.org/package=afex>.

Sommer, T., Hennies, N., Lewis, P. A., & Alink, A. (2022). The assimilation of novel information into schemata and its efficient consolidation. *The Journal of Neuroscience, 42* (30) 5916-5929. https://doi.org/10.1523/JNEUROSCI.2373-21.2022

Souza, C., Garrido, M. V., Horchak, O. V., & Carmo, J. C. (2022). Conceptual knowledge modulates memory recognition of common items: The selective role of item-typicality. *Memory & Cognition, 50*(1), 77–94. https://doi.org/10.3758/s13421-021-01213-x

Spalding, K. N., Jones, S. H., Duff, M. C., Tranel, D., & Warren, D. E. (2015). Investigating the neural correlates of schemas: Ventromedial prefrontal cortex is necessary for normal schematic influence on memory. *The Journal of Neuroscience*, *35*(47), 15746–15751. https://doi.org/10.1523/JNEUROSCI.2767-15.2015

Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: A neurobiological perspective. *Current Opinion in Neurobiology, 5*(2), 169–177. https://doi.org/10.1016/0959-4388(95)80023-9

Squire, L. R., Cohen, L., & Nadel, L. (1983). The medial temporal region and memory consolidation: A new hypothesis. In H. Weingartner & E. Parker (Eds.) *Memory consolidation* (S. 114–132). Erlbaum.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, *93*(24), 13515–13522. https://doi.org/10.1073/pnas.93.24.13515

Staresina, B. P., Gray, J. C., & Davachi, L. (2009). Event congruency enhances episodic memory encoding through semantic elaboration and relational binding. *Cerebral Cortex, 19*(5), 1198–1207. https://doi.org/10.1093/cercor/bhn165

Stites, M. C., Federmeier, K. D., & Christianson, K. (2016). Do morphemes matter when reading compound words with transposed letters? Evidence from eye-tracking and event-related potentials. *Language, Cognition and Neuroscience*, *31*(10), 1299–1319. https://doi.org/10.1080/23273798.2016.1212082

Stróżak, P., Abedzadeh, D., & Curran, T. (2016). Separating the FN400 and N400 potentials across recognition memory experiments. *Brain Research, 1635,* 41–60. https://doi.org/10.1016/j.brainres.2016.01.015

Sturm, V. E., Datta, S., Roy, A. R. K., Sible, I. J., Kosik, E. L., Veziris, C. R., Chow, T. E., Morris, N. A., Neuhaus, J., Kramer, J. H., Miller, B. L., Holley, S. R., & Keltner, D. (2020). Big smile, small self: Awe walks promote prosocial positive emotions in older adults. *Emotion 22*(5), 1044–1058. https://doi.org/10.1037/emo0000876

Sweegers, C. C. G., Coleman, G. A., van Poppel, E. A. M., Cox, R., & Talamini, L. M. (2015). Mental schemas hamper memory storage of goal-irrelevant information. *Frontiers in Human Neuroscience, 9.* https://doi.org/10.3389/fnhum.2015.00629

Takeuchi, T., Tamura, M., Tse, D., Kajii, Y., Fernández, G., & Morris, R. G. M. (2022). Brain region networks for the assimilation of new associative memory into a schema. *Molecular Brain, 15*(1), 24. https://doi.org/10.1186/s13041-022-00908-9

Teyler, T. J., & DiScenna, P. (1986). The hippocampal memory indexing theory. *Behavioral Neuroscience, 100*(2), 147–154.

Tibon, R., Cooper, E., & Greve, A. (2017). Does semantic congruency accelerate episodic encoding, or increase semantic elaboration? *The Journal of Neuroscience, 37*(19), 4861–4863. https://doi.org/10.1523/JNEUROSCI.0570-17.2017

Tibon, R., Greve, A., & Henson, R. (2018). The missing link? Testing a schema account of unitization. *Memory & Cognition*, *46*(7), 1023–1040. https://doi.org/10.3758/s13421-018-0819-3

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

Tierney N, Cook D, McBain M, Fay C (2021). naniar: Data Structures, Summaries, and Visualisations for Missing Data. R package version 0.6.1, <https://CRAN.R-project.org/package=naniar>.

Titchener, B. (1898). The postulates of a structural psychology. *The Philosophical Review, 7*(5), 449–465.

Tompary, A., Zhou, W., & Davachi, L. (2020). Schematic memories develop quickly, but are not expressed unless necessary. *Scientific Reports, 10*(1), 16968. https://doi.org/10.1038/s41598-020-73952-x

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

Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., Bito, H., & Morris, R. G. M. (2011). Schema-Dependent Gene Activation and Memory Encoding in Neocortex. Science, 333(6044), 891–895. https://doi.org/10.1126/science.1205274

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson, *Organization of memory.* (S. 381–403). Academic Press.

Tulving, E. (1983). *Elements of Episodic Memory*. Clarendon Press ; Oxford University Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology / Psychologie Canadienne*, *26*(1), 1–12. https://doi.org/10.1037/h0080017

Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annual Review of Psychology*, *53*(1), 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Urbach, T. P., & Kutas, M. (2002). The intractability of scaling scalp distributions to infer neuroelectric sources. *Psychophysiology, 39*(6), 791–808. https://doi.org/10.1111/1469-8986.3960791

Urbach, T. P., & Kutas, M. (2006). Interpreting event-related brain potential (ERP) distributions: Implications of baseline potentials and variability with application to amplitude normalization by vector scaling. *Biological Psychology*, *72*(3), 333–343. https://doi.org/10.1016/j.biopsycho.2005.11.012

van Kesteren, M. (2013). *Schemas in the brain: Influences of priorknowledge on learning, memory, and education.* Enschede, The Netherlands: Ipskamp Drukkers. van Kesteren, M. T. R., Beul, S. F., Takashima, A., Henson, R. N., Ruiter, D. J., & Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. *Neuropsychologia*, *51*(12), 2352–2359. https://doi.org/10.1016/j.neuropsychologia.2013.05.027

van Kesteren, M. T. R., Fernandez, G., Norris, D. G., & Hermans, E. J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences, 107*(16), 7550–7555. https://doi.org/10.1073/pnas.0914892107

van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., & Fernández, G. (2013). Consolidation Differentially Modulates Schema Effects on Memory for Items and Associations. *PLoS ONE, 8*(2), e56155. https://doi.org/10.1371/journal.pone.0056155

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

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83*(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015

Vilberg, K. L., Moosavi, R. F., & Rugg, M. D. (2006). The relationship between electrophysiological correlates of recollection and amount of information retrieved. *Brain Research, 1122*(1), 161–170. https://doi.org/10.1016/j.brainres.2006.09.023

von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychologische Forschung, 18,* 299–342.

Wagoner, B. (2013). Bartlett’s concept of schema in reconstruction. *Theory & Psychology, 23*(5), 553–575. https://doi.org/10.1177/0959354313500166

Wang, W. (2019). Cutting out the middleman: Separating attributional biases from memory deficits. *Behavioral and Brain Sciences, 42,* E302. https://doi.org/doi:10.1017/S0140525X1900195X

Wang, W.-C., Lazzara, M. M., Ranganath, C., Knight, R. T., & Yonelinas, A. P. (2010). The medial temporal lobe supports conceptual implicit memory. *Neuron, 68*(5), 835–842. https://doi.org/10.1016/j.neuron.2010.11.009

Wang, W.-C., Ranganath, C., & Yonelinas, A. P. (2014). Activity reductions in perirhinal cortex predict conceptual priming and familiarity-based recognition. *Neuropsychologia, 52*, 19–26. https://doi.org/10.1016/j.neuropsychologia.2013.10.006

Webb, C. E., Turney, I. C., & Dennis, N. A. (2016). What’s the gist? The influence of schemas on the neural correlates underlying true and false memories. *Neuropsychologia, 93*, 61–75. https://doi.org/10.1016/j.neuropsychologia.2016.09.023

Wechsler, D. (1955). *Wechsler Adult Intelligence Scale.* New York: Psychological Corporation.

Wentura, D., & Rothermund, K. (2014). Priming is not Priming is not Priming. *Social Cognition, 32*(Supplement), 47–67. https://doi.org/10.1521/soco.2014.32.supp.47

Westerman, D. L. (2008). Relative fluency and illusions of recognition memory. *Psychonomic Bulletin & Review, 15*(6), 1196–1200. https://doi.org/10.3758/PBR.15.6.1196

Westerman, D. L., Lloyd, M. E., & Miller, J. K. (2002). The attribution of perceptual fluency in recognition memory: The role of expectation. *Journal of Memory and Language, 47*(4), 607–617. https://doi.org/10.1016/S0749-596X(02)00022-0

Whittlesea, B. W. A., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don’t? A discrepancy-attribution account of feelings of familiarity. *Acta Psychologica, 98*(2–3), 141–165. https://doi.org/10.1016/S0001-6918(97)00040-1

Whittlesea, B. W. A., & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 547–565. https://doi.org/10.1037/0278-7393.26.3.547

Whittlesea, B. W. A., & Williams, L. D. (2001). The discrepancy-attribution hypothesis: I The heuristic basis of feelings and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 3–13. https://doi.org/10.1037/0278-7393.27.1.3

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H. (2019). “Welcome to the tidyverse.” *Journal of Open Source Software, 4* (43), 1686. doi:10.21105/joss.01686.

Wickham, H., François, R., Henry, L., Müller, K. (2022). dplyr: A Grammar of Data Manipulation. R package version 1.0.10, <https://CRAN.R-project.org/package=dplyr>.

Wiegand, I., Bader, R., & Mecklinger, A. (2010). Multiple ways to the prior occurrence of an event: An electrophysiological dissociation of experimental and conceptually driven familiarity in recognition memory. *Brain Research, 1360*, 106–118. https://doi.org/10.1016/j.brainres.2010.08.089

Wilding, E. L. (2006). The practice of rescaling scalp-recorded event-related potentials. *Biological Psychology, 72*(3), 325–332. https://doi.org/10.1016/j.biopsycho.2005.12.002

Winocur, G., & Moscovitch, M. (2011). Memory Transformation and Systems Consolidation. *Journal of the International Neuropsychological Society, 17*(05), 766–780. https://doi.org/10.1017/S1355617711000683

Winocur, G., Moscovitch, M., & Bontempi, B. (2010). Memory formation and long-term retention in humans and animals: Convergence towards a transformation account of hippocampal–neocortical interactions. *Neuropsychologia, 48*(8), 2339–2356. https://doi.org/10.1016/j.neuropsychologia.2010.04.016

Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity.

Brain Research, 1100(1), 125–135. https://doi.org/10.1016/j.brainres.2006.05.019

Woollams, A. M., Taylor, J. R., Karayanidis, F., & Henson, R. N. (2008). Event-related Potentials Associated with Masked Priming of Test Cues Reveal Multiple Potential Contributions to Recognition Memory. *Journal of Cognitive Neuroscience, 20*(6), 1114–1129. https://doi.org/10.1162/jocn.2008.20076

Yacoby, A., Reggev, N., & Maril, A. (2021). Examining the transition of novel information toward familiarity. *Neuropsychologia, 161,* 107993. https://doi.org/10.1016/j.neuropsychologia.2021.107993

Yagoubi, R. E., Chiarelli, V., Mondini, S., Perrone, G., Danieli, M., & Semenza, C. (2008). Neural correlates of Italian nominal compounds and potential impact of headedness effect: An ERP study. *Cognitive Neuropsychology*, *25*(4), 559–581. https://doi.org/10.1080/02643290801900941

Yang, X., Mao, X., Han, M., Li, X., & Guo, C. (2019). Electrophysiological evidence of encoding in self-referential effect: *NeuroReport*, *30*(13), 901–907. https://doi.org/10.1097/WNR.0000000000001303

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, *34*(10), 515–525. https://doi.org/10.1016/j.tins.2011.06.006

Yonelinas, A. P. (1994). Receiver-Operating Characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1341–1354.

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

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

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

Yonelinas, A. P., Hopfinger, J. B., Buonocore, M. H., Kroll, N. E. A., & Baynes, K. (2001). Hippocampal, parahippocampal and occipital-temporal contributions to associative and item recognition memory: An fMRI study: *Neuroreport, 12*(2), 359–363. https://doi.org/10.1097/00001756-200102120-00035

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

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychological Bulletin, 133*(5), 800–832. https://doi.org/10.1037/0033-2909.133.5.800

Züst, M. A., Ruch, S., Wiest, R., & Henke, K. (2019). Implicit vocabulary learning during sleep is bound to slow-wave peaks. *Current Biology, 29*(4), 541-553.e7. https://doi.org/10.1016/j.cub.2018.12.038

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes*, *9*(3), 341–368. https://doi.org/10.1080/01690969408402123