Effects of similarity in working memory

Contrasting mutual facilitation, a retrieval account, and inhibitory processes

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

der Fakultät HW

Bereich Empirische Humanwissenschaften

der Universität des Saarlandes

vorgelegt von

Demian Scherer

aus Ottweiler

Saarbrücken, 2020

Dekan:

Univ.-Prof. Dr. Stefan Strohmeier

Berichterstatter:

Prof. Dr. Dirk Wentura, Universität des Saarlandes

Prof. Dr. Hubert Zimmer, Universität des Saarlandes

Tag der Disputation: 16.12.2019

Acknowledgment

Many people supported me in achieving this thesis. I want to thank:

- my supervisor and mentor Dirk Wentura who always had an open ear and open door for me. I want to thank him for his advice, the support, and the appreciation.

- Melanie Schmitz for interesting discussions and for advice during the development of this work.

- my colleagues Alex, Benjamin, Benedikt, as well as Andrea, Charlotte, Katrin, Michaela, and Timea who brightened up many of my days during countless Mensa breaks.

- the student research assistants Jonas and Nancy and the interns Franziska, Daniel and again Jonas who collected many data for this thesis and other studies.

- Glen Forester for checking the spelling and giving advice on grammar issues.

- my friends who always believed in me.

- my family, especially my mother for her support in many areas. I also want to deeply thank my father who early on paved the way for my academic career and supported me in all areas.

- Elisabeth for her support and for always being there.

Author note

Part of this work is based on manuscripts that are either submitted or in preparation. Specifically, Chapter 2 includes Experiments 1a-b, which are based on Scherer & Wentura, 2018. Experiment 2a-c is covered in a manuscript which is currently in preparation for submission. For these manuscripts, I am the first author. In order to warrant a smooth reading, the respective parts are not marked; in addition, I consistently employ the term "we" instead of "I" throughout the whole thesis. Experiment 1a was also part of my master thesis. However, it is included in this thesis to provide a more thorough overview.

Contents

List of Tables

List of Figures

List of Abbreviations

%......... percent
€ .... Euro
α .... rate of type I error
β .... rate of type II error
η_p^2......... effect size for ANOVAs with repeated measures (partial η^2)
a .... threshold separation (drift diffusion model parameter)
a_{low}......... threshold speration parameter for the low threshold
a_{high} .... threshold speration parameter for the high threshold
ANOVA .... analysis of variance
d' .... measure of sensitivity used in signal detection theory
df .... degrees of freedom
d_z .... effect size for paired t-tests
EEG .... electroencephalography
e.g......... for example
ERP .... event-related potential
F......... test statistic from F-distribution
KDEF .... Karolinska Directed Emotional Faces databank
KS......... Kolmogorov-Smirnov
M .... mean
ms......... millisecond
p. .... page
p .... p value (the observed significance level)
pp .... pages
s .... second
S-R .... stimulus-response
S-S......... stimulus-stimulus
SOA......... stimulus onset asynchrony
t .... test statistic from Student's t-distribution
t_0......... nondecision time (drift diffusion model parameter)
vs .... versus
v......... drift rate (drift diffusion model parameter)
z .... starting-point (drift diffusion model parameter)
zr .... relative starting point (drift diffusion model parameter)

Zusammenfassung

Gibt es automatische Prozesse, die uns helfen, ähnliche Konzepte für kurze Zeit besser zu behalten als unähnliche Konzepte? Einige ältere Studien, welche den Effekt von Ähnlichkeit auf die Arbeitsgedächtnisleistung untersucht haben, zeigen widersprüchliche Befunde. Zudem untersuchen diese Studien nicht zwangsläufig automatische Prozesse, die durch konzeptuelle oder semantische Ähnlichkeit verursacht werden. *Spreading activation*-Theorien und einige andere Theorien, die versuchen Primingeffekte zu erklären, können vielversprechende Ansatzpunkte für das Verständnis automatischer Effekte von Ähnlichkeit auf die Gedächtnisleistung liefern. Zum Beispiel kann, basierend auf den Annahmen von *spreading activation*-Theorien, vorhergesagt werden, dass ähnliche Elemente im Gegensatz zu unähnlichen Elementen leichter gleichzeitig im Arbeitsgedächtnis gehalten werden können, da sie wechselseitig ihre Aktivation aufrechterhalten.

Vor diesem Hintergrund wurde mit Experiment 1a die Grundlage dafür geschaffen, Evidenz für eine parallele Aktivation und eine automatische wechselseitige Aufrechterhaltung von Prime und Target in einer semantischen Primingaufgabe zu liefern. Darauf aufbauend liefert Experiment 1b Evidenz für das Konzept der parallelen Aktivierung und der wechselseitigen Aufrechterhaltung im Priming. Darüber hinaus wurde in Experiment 1b eine Gedächtniskomponente in die Primingaufgabe eingebaut, indem die perzeptuelle Identifikationsaufgabe (die auch im Experiment 1a verwendet wurde) mit einer *post cue*-Aufgabe kombiniert wurde. Die folgenden Studien zielen darauf ab, Evidenz für die wechselseitige Aufrechterhaltung ähnlicher Konzepte im Arbeitsgedächtnis zu sammeln. Zu diesem Zweck wurde eine klassische Arbeitsgedächtnisaufgabe eingesetzt: die *change detection*-Aufgabe. In vier Experimenten (Experiment 2a-d) dienten emotionale Gesichter (wütende und fröhliche Gesichter) als zu erinnernde Stimuli. Dies ermöglichte die Untersuchung des Effekts evaluativer Kongruenz, die als eine spezifische Art konzeptueller Ähnlichkeit aufgefasst werden kann. Obwohl die Ergebnisse der einzelnen Experimente (Experiment 2a-d) uneindeutig erschienen, zeigte eine Gesamtanalyse eine bessere Gedächtnisleistung in

Durchgängen mit evaluativ kongruenten verglichen mit evaluativ inkongruenten Gesichtern. Dieses Ergebnis stimmt mit der Annahme einer wechselseitigen Aufrechterhaltung aufgrund konzeptueller Ähnlichkeit überein. Die wechselseitige Aufrechterhaltung ist dabei ein Prozess, der vermutlich im Arbeitsgedächtnis zu verorten ist. In einer Reihe ähnlicher Experimente (Experiment 3a-c) wurde eine andere Variante des *change detection*-Paradigmas verwendet (mit zwei Elementen pro Display und wechselnden Positionen von der Enkodierung zum Test) um wiederum den Effekt evaluativer Kongruenz zu untersuchen. Mit dieser unterschiedlichen Prozedur zeigte sich jedoch insgesamt eine bessere Leistung in inkongruenten Durchgängen, gemessen mit *d‘* sowie ein entsprechender numerischer Effekt im Driftdiffusionsparameter *v*. Auf den ersten Blick widerspricht dieser Befund der Annahme einer wechselseitigen Aufrechterhaltung evaluativ kongruenter oder konzeptuell ähnlicher Konzepte. Es gilt jedoch zu beachten, dass der Wechsel der Positionen der Items vom Enkodierdisplay zum Testdisplay in Experiment 3a-c einen entgegenwirkenden Prozess verursacht haben könnte. In inkongruenten Durchgängen könnten die Teilnehmer die Emotionen der beiden Gesichter genutzt haben, um zu erschließen, welches Gesicht im Testdisplay mit welchem Gesicht aus dem Lerndisplay verglichen werden muss. Diese vermutlich höchst effiziente Strategie kann jedoch nicht in kongruenten Durchgängen eingesetzt werden. Weiterhin sollte diese Strategie ohne Nutzen sein, wenn zwei Elemente pro Display ohne wechselnde Positionen vom Lern- zum Testdisplay verwendet werden. Diese Annahme wurde in Experiment 4 getestet. In dieser letzten Studie wurde ein Gedächtnisvorteil aufgrund evaluativer Kongruenz beobachtet, der sich mithilfe der Annahme einer wechselseitigen Aufrechterhaltung erklären lässt. Das Design von Experiment 4 beinhaltete jedoch einen weiteren Faktor: Es wurde variiert, ob die ursprünglich gelernte Konfiguration im Testdisplay durch eine aufgabenirrelevante Veränderung zerstört wird oder nicht. Der Effekt einer solchen irrelevanten Veränderung wirkte sich jedoch nicht stärker auf die Leistung in kongruenten Durchgängen als auf die Leistung in inkongruenten Durchgängen aus. Somit zeigte sich keine direkte Evidenz für eine Alternativerklärung (die auf einer *compound-cue*-Theorie basiert, genauer dem *retrieval account* von Whittlesea und Jabocy, 1990). Diese Alternativerklärung unterscheidet sich grundlegend von der Annahme einer wechselseitigen

Aufrechterhaltung auf der Basis von Aktivationsausbreitungsprozessen. Daher können die hier berichteten Effekte am einfachsten entweder durch eine wechselseitige Aufrechterhaltung ähnlicher Konzepte oder durch die Verwendung der Valenz der Gesichter als Abrufschlüssel (Experiment 3a-c) erklärt werden. Mögliche Alternativerklärungen durch *compound-cue*-Theorien, Modifikationen aktueller parallelverteilter Modelle oder durch die Modifikation von Theorien, die in erster Linie Effekte perzeptueller Ähnlichkeit erklären, werden diskutiert.

Abstract

Are there automatic processes that help us to memorize similar concepts more easily than dissimilar concepts for a short time? Early studies investigating the effects of similarity on working memory[1] performance revealed diverging effects. Furthermore, these studies do not necessarily address automatic processes that are induced by conceptual or semantic similarity. Theories that attempt to explain priming effects, especially spreading activation theories, may also provide a promising framework for understanding the automatic influences of similarity on task performance. Based on the assumptions of spreading activation theories, it can be predicted that similar items are more easily maintained concurrently in working memory than dissimilar items because they mutually facilitate each other's activation.

As a basis, in Experiment 1a we paved the way to provide evidence for parallel activation and mutual facilitation of prime and target in a semantic priming task that can be assumed to measure automatic priming effects. Based on that, Experiment 1b provides evidence for the notion of parallel activation and mutual facilitation in priming. In addition, Experiment 1b introduces a memory component into the priming task by combining the perceptual identification task, which was also used in Experiment 1a, with a post-cue task. In the next step, we aimed for providing evidence for mutual facilitation processes in working memory. Therefore, we used a classical working memory paradigm: the change detection task. In four experiments (Experiment 2a-2d), emotional faces (i.e., angry and happy faces) served as to-be-remembered stimuli in order to enable the investigation of the evaluative congruency effect because evaluative congruency can be considered to be a specific type of conceptual similarity. Although the results of the individual experiments (Experiment 2a-2d) were heterogeneous, an overall analysis revealed better performance in trials with evaluatively congruent

[1] In this thesis, the term working memory is used for a memory system with a limited capacity that is used to temporarily hold information in an active state. We assume that information in working memory can be activated long-term memory representations (e.g. Cowan, 1999). Further, we use the terms working memory and short-term memory as interchangeable (as for example Luck & Vogel, 1997), knowing that the term working memory is sometimes only used when information is not only maintained but also manipulated (Engle, 2002).

compared to evaluatively incongruent content. This finding is in accordance with the assumption of a mutual facilitation due to conceptual similarity that occurs in working memory. In a series of similar experiments (Experiment 3a-3c), in which another variant of the change detection paradigm was implemented (with set size two and changing locations from study to test), again the effect of evaluative congruency was investigated. However, using a different procedure, there was overall (Experiment 3a-c) better performance in incongruent trials measured with *d'* and a numerical effect in the drift-diffusion parameter *v*. At first glance, this finding stands in contrast to the assumption of a mutual facilitation of evaluatively or conceptually similar items. Notwithstanding, the change of locations from study to test used in Experiment 3a-c might have introduced a counteracting process. In incongruent trials, participants might have used the emotions of the two faces to infer which test face has to be compared with which studied face. This potentially efficient strategy cannot be used in congruent trials. The strategy is however of no use when set size two without locational changes from study to test is used. This assumption was tested in Experiment 4. In this final study, an overall memory benefit due to evaluative congruency was observed, which is in line with the assumption of a mutual facilitation. The design of Experiment 4 included an additional factor manipulating whether the studied arrangement was destroyed at test by an irrelevant change or not. The effect of the irrelevant change, however, did not affect performance in congruent trials more than in incongruent trials. Therefore, there is no direct evidence for an alternative explanation (based on a compound-cue theory, namely the retrieval account by Whittlesea and Jacoby, 1990) that is different to the assumption of a mutual facilitation based on spreading activation processes. Hence, the effects of the reported studies can most parsimoniously be explained either by mutual facilitation of similar concepts or by the use of valence as a retrieval cue in Experiment 3a-c. Potential alternative explanations by compound-cue theories, modifications of parallel distributed processing models or theories inspired by explanations of effects of perceptual similarity are discussed.

Preface

Imagine a friend listing brands of luxury cars, which you would like to drive: *Porsche*, *Ferrari*, *Audi*, *BMW*, *Mercedes*, *Bentley,* and *Bugatti*. Can you remember these names easily? Now imagine another friend listing completely different concepts: *ball*, *Audi*, *computer*, *t-shirt*, *banana*, *suitcase*, and *letter*. Which of these lists would be easier to reproduce? The answer seems simple: Intuition suggests that the first list containing several related concepts should be retrieved more easily.

But is there an automatic effect of relatedness contributing to this plausible phenomenon? And what precisely happens when we maintain pieces of related information in our working memory? In other words: Are there automatic effects of relatedness on working memory performance? There is a long-lasting research tradition analyzing the effect of relatedness in memory, using semantically related and unrelated words as stimuli. In these studies, participants have to remember words in short-term or working memory tasks. In many of these experiments, participants performed better when the to-be-remembered words were related (Cowles, Garnham, & Simner, 2010, Experiment 2; Goh & Goh, 2006; Murdock & Vom Saal, 1967; Oberauer, 2009b; Poirier & Saint-Aubin, 1995; Saint-Aubin, Ouellette, & Poirier, 2005). In other experiments, no significant effects or mixed results were observed (Baddeley & Levy, 1971; Cowles et al., 2010, Experiment 1; van der Lely & Howard, 1993). However, there are also studies in which performance declined when the to-be-remembered words were related (Baddeley, 1966; Dale & Gregory, 1966). In any case, the question arises whether a potential advantage or disadvantage in performance due to semantic relatedness is (merely) based on strategic processes or whether these studies can provide insight into automatic processes of relatedness on working memory performance. Think back to the example before: For *Porsche*, *Ferrari*, *Audi*, *BMW*, *Mercedes*, *Bentley,* and *Bugatti* you can simply generate the common category, in this case: *luxury car brands.* When you are instructed to retrieve the learned words, you can benefit from generating specific exemplars of luxury car brands using the category that you might remember as a cue. Indeed, this might lead to specific false responses; however, only a few studies have inspected false responses for the generation of not learned category members (e.g.

Crowder, 1979; Poirier & Saint-Aubin, 1995, Experiment 3).[2] Nonetheless, controlling for false responses can only rule out a subset of strategic processes. Crowder (1979), as well as Poirier and Saint-Aubin (1995), assume that their effects might also be caused by more sophisticated guessing strategies (e.g., using a combination of knowledge about the category type and partial memory information of the items). These guessing strategies may not be regarded as automatic processes, which are often defined by one or more separate features such as "unintentional, uncontrolled/uncontrollable, goal independent, autonomous, purely stimulus driven, unconscious, efficient, and fast" (Moors & De Houwer, 2006, p. 297). Nevertheless, their studies offer a good starting point: They show which guessing strategies should be circumvented when analyzing automatic processes caused by semantic similarity. Furthermore, assuming for a moment that automatic processes contributed to the effects in these studies, one might speculate that relatedness could lead to enhanced performance (which was observed in many studies, e.g. Cowles et al., 2010, Experiment 2; Goh & Goh, 2006; Murdock & Vom Saal, 1967; Oberauer, 2009b; Poirier & Saint-Aubin, 1995; Saint-Aubin et al., 2005).

There are studies that could provide a foundation for an analysis of automatic processes of semantic relatedness. They might provide hypotheses about the influence of relatedness on working memory performance. These studies compare working memory performance for stimuli from different categories with working memory performance for stimuli stemming from a single category (M. A. Cohen, Konkle, Rhee, Nakayama, & Alvarez, 2014; Jiang, Remington, Asaad, Lee, & Mikkalson, 2016). For example, memory for faces and scenes can be compared with memory for only faces or only scenes. While most studies manipulating relatedness in word lists observed beneficial effects of relatedness on working memory performance, these studies using different semantic categories mostly observed an advantage for dissimilar stimuli. This finding does not necessarily indicate an automatic effect of *semantic* overlap. Besides the

[2] Please note that the Deese-Roediger-McDermott (DRM) paradigm is a specific case in which errors of this kind arise. In their tasks, subjects learn lists of words like *bed*, *rest*, *awake,* and so on, which are associates of a nonpresented word (in this case *sleep*). In a following free recall, usually a relatively high false recall rate for the nonpresented associate is observed. Further, in subsequently tested recognition, the false alarm rate for these words is at a comparable level as the hit rate for the presented words (e.g., Roediger & McDermott, 1995).

manipulation of semantic relatedness by using displays depicting faces and/or scenes, in these studies, there are also remarkable differences in the *perceptual* overlap between the "related" and the "unrelated" condition. Accordingly, the effect could also be explained by the tremendous perceptual differences between the distinct stimulus groups (Jackson, Linden, Roberts, Kriegeskorte, & Haenschel, 2015). Interestingly, for studies in the visual domain, when perceptual (and conceptual) similarity was manipulated within a category and using complex stimuli, beneficial effects of similarity were observed (Jiang, Lee, Asaad, & Remington, 2016). However, these studies do not provide a satisfying, clear and unambiguous theoretical background for providing insight into the automatic effects of semantic relatedness on working memory performance. Rather, they again offer valuable insight into what should be controlled for when more pure and automatic effects of semantic relatedness on working memory performance are the focus.

Indeed, there is an area of research that can provide a solid foundation for an analysis of automatic processes, namely the research on semantic and evaluative priming. However, priming theories and evidence for automatic semantic and evaluative priming seem at first glance unrelated to working memory research. Nevertheless, in the current thesis exactly this starting point was chosen: Theories and procedures originating from the research on semantic and evaluative priming were adapted and employed to the area of working memory research. Thereby, ideas on how to predict and how to measure effects of semantic and evaluative congruency in working memory were generated. Accordingly, the following first chapter, which addresses the theoretical background of this thesis, mainly focuses on theories that are utilized to explain semantic and evaluative priming and their application to working memory research. In this chapter, we aim to convince the reader that semantic and evaluative priming on the one hand, and working memory on the other, are more related than might be assumed at first glance.

1 The link between priming and working memory research

Focusing on automatic processes (in the sense of uncontrollability and unintentionality), a well suited theoretical framework for effects of relatedness on working memory performance are models and theories from a seemingly unrelated but highly influential area of psychological research: the literature on semantic and evaluative priming. To explain why research on semantic and evaluative priming can be applied to the field of working memory research, it is worthwhile to take a closer look at the definition of semantic priming.

Semantic priming is frequently defined as an "improvement in speed or accuracy to respond to a [target] stimulus when it is preceded by a semantically related or associated [prime] stimulus relative to when it is preceded by a semantically unrelated or unassociated stimulus (e.g., *cat-dog vs. table-dog* ...)" (McNamara, 2013, p. 455). Typically, in semantic priming studies, the target is categorized as a word (versus non-word) or it has to be named as soon as possible; the prime is presented very briefly and it *precedes* the target (for reviews, see McNamara, 2005, 2013; Neely, 1991). At first glance, this definition of semantic priming, as well as the label "prime" for the preceding stimulus, alludes to the usual temporal order of the stimuli. Therefore, with the focus on the sequence of presentation one might speculate that concepts also must be active one after the other (Masson, 1991, 1995). Accordingly, semantic priming theories might differ with respect to the question: Does the theory assume a temporal order of prime and target activation, and if so, can the prime still be active after processing the target, or not? In other words, can a priming theory allow for a *concurrent* activation of the prime and the target concept, or does the theory rely on the more usual sequence of a prime that is followed by a target? In principle, priming theories that allow for concurrent activation of prime and target allow for the parallel activation of two or more concepts. Thus, if a priming theory allows for parallel activation of prime and target, the theory can easily make predictions about automatic processes of relatedness on performance in working memory tasks, because working memory is a system in which two or more concepts can be active simultaneously (Brady, Konkle, & Alvarez, 2011; Cowan, 1988, 1995, 1999, 2010; Luck & Vogel, 1997, 2013). In contrast, priming theories that rely on the temporal order of prime and target activation would make a linkage between priming and working memory phenomena impossible or at least extremely difficult to implement (for this argument, see also Scherer & Wentura, 2018; Schmitz & Wentura, 2012; Schmitz, Wentura, & Brinkmann, 2014). Because the main focus of this thesis is on effects of relatedness in working memory, we will initially focus on theories that clearly allow for parallel activation of prime and target. Theories that might need modification to be applied to working memory paradigms (e.g. the model by Masson, 1991, 1995) are discussed later.

Evaluative priming can be regarded as a specific kind of semantic priming in which the overlap in the evaluative component defines the similarity between concepts. Evaluative priming refers to the finding that responses to an evaluatively valenced target (e.g., the word *love*) are initiated faster and/or they are more accurate, when an evaluatively congruent prime precedes the target (e.g., the word *sunshine*) rather than an evaluatively incongruent prime (e.g., the word *death*; see Klauer & Musch, 2013). Typically, the effect is found with the evaluative decision task. That is, the target has to be categorized as positive or negative such that an evaluative priming effect can be explained by response tendencies triggered by the prime: When the prime triggers the same response as required by the target there can be facilitation; when the prime is associated with a divergent response tendency there is response conflict. Therefore, priming effects of this kind are also labeled *stimulus-response (S-R) based evaluative priming effects* (see De Houwer, 2003; Schmitz & Wentura, 2012; Schmitz et al., 2014). However, there is also evidence for evaluative priming effects that cannot be based on response processes, for example, if a priming effect is found with the naming task (i.e., the target has to be named; see Herring et al., 2013, for meta-analytical support, but see Klauer, Becker, & Spruyt, 2016; Klauer & Musch, 2001). These non-response based evaluative priming effects are mainly explained by the same theories as classical semantic priming. This type of priming effect is also termed *stimulus-stimulus (S-S)-based evaluative priming effect* (see De Houwer, 2003; Schmitz & Wentura, 2012; Schmitz et al., 2014). Basically, stimulus-stimulus based evaluative priming effects can be explained by the same theoretical frameworks as other semantic priming effects that are described later (for example spreading activation theories can explain S-S-based evaluative priming, when valence nodes are assumed as in the model by Bower, 1991; see Schmitz, 2012 and Spruyt, Hermans, Houwer, & Eelen, 2002 for a discussion of this assumption). Thus, when there are beneficial effects of semantic relatedness on working memory performance that can be explained by the same theories as semantic priming effects, and when semantic priming and S-S-based evaluative priming can be explained by the same theories, there should also be beneficial effects of evaluative congruency on working memory performance. However, it should be noted that the effects for S-S based evaluative priming seem to be less pronounced than the effects for stronger semantic associates as stimuli (for a meta-analysis on evaluative priming see Herring et al., 2013).

In the following, theories are considered that can make predictions about the effects of semantic similarity (and the subtype of evaluative congruency) on working memory performance. While the first of the mentioned theories have their roots in priming research, the later partially address memory more directly. Here, only priming theories are described that account for a parallel activation of the prime and the target concept and thus more generally the parallel activation of several concepts. Therefore, these theories can directly be linked to working memory phenomena. A discussion of other priming theories that rely on a sequential activation of prime and target concept will be postponed to the discussion of priming studies in this thesis (Experiment 1a and 1b) and the General Discussion. After a description of the central theories, relevant evidence will be considered, followed by an overview of the experiments conducted for this thesis.

1.1 Priming theories

1.1.1 The Three-Process Model

A direct link between priming phenomena and the working memory system was built in the context of the *three-process model* by Schmitz and Wentura (Schmitz & Wentura, 2012; see also Schmitz et al., 2014). Their model was designed to explain evaluative priming effects, and they propose three processes that are potentially relevant in evaluative priming tasks. However, these three processes, which were derived from the theoretical account of Wentura and Rothermund (2003), can in principle also take place when, not evaluative congruency, but a different kind of similarity causes priming.

The first process is parallel activation of prime and target representations. From a certain perspective, this point might seem obvious because it is a necessary precondition to explain the S-R-based evaluative priming effects that were described above. The second process is the mutual facilitation of evaluatively congruent concepts (or when applied more generally: semantically similar concepts). This process can be subdivided into two parts that will be labeled in the following as *proactive* and *retroactive* priming. Proactive priming is understood in the following as the facilitation of target processing by a semantically related (e.g., evaluatively congruent) prime. This proactive priming is usually understood as the standard case of "priming". However, the target can further help to maintain the activation of an evaluatively congruent and/or semantically related prime. While Wentura and colleagues provide evidence for this assumption in evaluative priming (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008), in the research of semantic priming, the effect of related targets on the prime are often described with the label *retroactive priming* (Briand, den Heyer, & Dannenbring, 1988; Dark, 1988; see section 1.4.1). To conclude, the second process by Schmitz and Wentura (2012) describes a facilitative effect in the case of relatedness/congruency in two directions: On the one hand, the prime activation facilitates target encoding and, on the other hand, the target activation helps to maintain the prime. The third process concerns response processes in priming. When both prime and target are associated with a response, these responses can either be compatible or incompatible. Therefore, two potential influences on performance can arise: response facilitation and response conflict. Please note, however, that in several standard priming paradigms in which the prime is not directly associated with a response, neither of the two potential influences are of importance. The same is true for potential applications of the three-process model to working memory tasks in which response conflict mostly does not influence results.

For the current purpose, the most interesting point about the three-process model is that the authors specifically suggest that the mutual facilitation of simultaneously activated related concepts, for which they provide evidence utilizing priming and a flanker paradigm (Schmitz & Wentura, 2012; Schmitz et al., 2014; see section 1.4.1), might arise from working memory. The assumption is plausible, especially because of the first process of the model. In working memory, two or more concepts can be held active at the same time (Brady et al., 2011; Cowan, 1988, 1995, 1999, 2010; Luck & Vogel, 1997, 2013), and in priming, the first process assumed by Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014) predicts exactly the same: a parallel activation of (at least) two concepts: the prime and the target. Therefore, when parallel activation can be linked to working memory, mutual facilitation of related prime-target pairs, which is the second process of the three-process model, could also be located in working memory. As outlined above, the three-process model can be easily applied to both evaluative congruency and semantic relatedness. Accordingly, the central claim of this thesis can be derived from the theory by Schmitz and Wentura (2012), which is: Semantically related concepts can mutually facilitate each other's activation in working memory. This should lead to better performance for semantically related items (as well as evaluatively congruent items) compared to semantically unrelated items (or evaluatively incongruent items) in both priming and working memory tasks.

Of course, one might reject the linkage of priming and working memory research as suggested here and by Schmitz and Wentura (2012; Schmitz et al., 2014). Criticism in this way can be based on the notion that priming is a phenomenon of long-term memory, whereas working memory deals with the short-term storage of items. Nevertheless, one of the most influential models of working memory, the embedded-processes model of working memory by Cowan (1988, 1995, 1999), conceptualizes the contents of working memory as a subset of the currently activated representations of long-term memory (see also Oberauer, 2002, 2006, 2009a; Oberauer & Lange, 2009; Oberauer, Souza, Druey, & Gade, 2013). Given this framing, the question arises of how we can model a long-term memory system that allows for parallel activation of concepts that mutually facilitate each other. This also provides a new perspective on spreading activation theories that are discussed in the next section. A more detailed description of these thoughts is given when the theory and its corresponding evidence by Davelaar and colleagues (Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Davelaar, Haarmann, Goshen-Gottstein, & Usher, 2006; see also Haarmann & Usher, 2001; Usher & Cohen, 1999) are considered (see section 1.2.1).

While the three-process model can make the clear and precise prediction that there is a beneficial process that leads to performance enhancement due to some kind of semantic similarity, it does not state how, precisely, the assumed mutual facilitation operates. As the authors suggest, *spreading activation theories* do provide concrete mechanisms that are well suited to explain priming and – most importantly – a mutual facilitation of related concepts. Therefore, also from the perspective of working memory research focusing on effects of semantic similarity, it is worthwhile to take a closer look at spreading activation theories and to consider whether they can account for a parallel activation of several concepts. As already anticipated, the discussion of priming theories that rely on a successive activation of prime and target concept is postponed to the Discussion on Experiment 1 and the General Discussion.

1.1.2 Spreading activation theories

The simplest and earliest accounts of semantic priming are the spreading activation theories (J. R. Anderson, 1976, 1983, 1993; Collins & Loftus, 1975). There, within a conceptual network, single nodes represent single concepts. In other words, spreading activation theories assume localist representations of concepts. Once the activation of a concept reaches a certain threshold, activation spreads along links to related concepts, and activation decays the farther the distance is. Thus, the degree to which activation will spread from an active concept to a second concept depends on their relatedness. For example, once the concept *cat* is activated, activation will spread to the related concept *dog*, whereas activation from the concept *cat* will not reach the distant part of the network in which the node for the concept *truck* is located. Consequently, the spread of activation preactivates related but not unrelated concepts. Semantic priming effects are explained by this initial activation of related but not unrelated concepts. A related concept will have greater preactivation than an unrelated concept, and will therefore be more likely to fully activate, and will be able to do so at a faster rate. When it is assumed that positive and negative concepts are linked to corresponding valence nodes for positive and negative valence (Bower, 1991), spreading activation theories can predict S-S-based evaluative priming effects.

But how precisely can we apply spreading activation theories to predict effects of relatedness in working memory paradigms? Suppose that two representations both reach the threshold to trigger a spread of activation because they are in an active state. In this case, different predictions arise for related and unrelated concepts when they are concurrently active. For related concepts, activation would spread back and forth between the two similar concepts. This would lead to an automatic mutual maintenance of their activation. In contrast, for two unrelated concepts, the spread of activation from one concept would not reach the distant part of the networks in which the other, unrelated concept is active. Therefore, not being automatically maintained by a mutual facilitation due to a spread of activation, unrelated concepts would more likely be "forgotten". Thus, spreading activation theories provide an explanation for the second process that was assumed in the three-process model by Schmitz and Wentura, which is the mutual facilitation of related concepts (Schmitz & Wentura, 2012; Schmitz et al., 2014). Assuming that the contents of working memory are currently activated concepts (that are in the focus of attention and awareness) as described in the embedded-processes model of working memory by Cowan (1988, 1995, 1999 see also Oberauer, 2002, 2006, 2009b; Oberauer & Lange, 2009; Oberauer et al., 2013 for similar assumptions), the mutual facilitation would take place in the working memory system.

Please note again that while spreading activation (with concepts being represented by single nodes) can easily be applied to working memory tasks, other theories of semantic priming focus more on the temporal succession of prime and target to explain priming effects (for an overview, see McNamara, 2005, 2013; Neely, 1991). The latter theories (e.g. Masson, 1991, 1995) possibly need some additional assumptions to be applied to working memory tasks or to account for evidence of parallel activation and mutual facilitation in priming.

Mutual facilitation in the theory by Collins and Loftus (1975). The assumptions of parallel activation of several concepts and mutual facilitation between them when they are related can be explained by the spreading activation theory by Collins and Loftus (1975) if the theory is marginally modified. That is because, in the original theory, it was assumed that "activation can only start out at one node at a time" (p. 411). When this unnecessary assumption is rejected, their model would predict that activation spreads back and forth between related activated concepts, leading to mutual facilitation. This minor modification does not affect central features of their theory. Thus, the theory can provide concrete mechanisms that can explain the mutual facilitation for related concepts (or evaluatively congruent concepts) that was assumed by Schmitz and Wentura for priming (Schmitz & Wentura, 2012; Schmitz et al., 2014). As outlined, this mutual facilitation can take place in working memory, especially when working memory is regarded as part of the activated long-term memory (Cowan, 1988, 1995, 1999 see also Oberauer, 2002, 2006, 2009b; Oberauer & Lange, 2009; Oberauer et al., 2013 for similar assumptions). Thus, the framework by Collins and Loftus (1975) can predict better performance for semantically related items (e.g. evaluatively congruent items) compared to unrelated items in working memory tasks due to mutual facilitation of related concepts.

Mutual facilitation in ACT*. Interestingly, the ACT* Model by Anderson (J. R. Anderson, 1983 see also J. R. Anderson, 1976, 1993 for other versions of the model) does not make the questionable assumption by Collins & Loftus (1975) that activation can only start out at one node at a time. McNamara (2013) posits that a central difference between the spreading activation theory ACT* by J. R. Anderson (1983), and Collins and Loftus' (1975) spreading activation theory, is that in ACT* prime and target *must* both be sources of activation, and therefore objects of attention, to produce priming effects. Therefore, it can be stated that prime and target concepts that are active in working memory should produce priming. Whether or not the theory by Anderson implies that concepts that are only weakly active and that did not enter working memory also cause priming is less relevant for our purpose. What is more important is that once a concept has entered the working memory system, it is clear that the concept is – according to the assumptions by the ACT* model and also according to the assumption by Collins and Loftus (1975) – most likely (if not certainly) activated enough to cause a spread of activation. Further, in ACT*, priming does not occur *despite* the prime being still in an active state, it takes place *because* the prime is still active when the target appears. In contrast, in the theory by Collins and Loftus (1975), the target can also be preactivated when the prime is no longer active.

To conclude, spreading activation theories, in general, are compatible with a parallel activation of several concepts and the notion of a mutual facilitation in the case of relatedness. Therefore, they can be applied to predict beneficial effects of semantic similarity in working memory that are caused by mutual facilitation. What is more, with ACT* there is also a theory that assumes that parallel activation of prime and target (potentially in working memory) is *mandatory* to obtain priming effects. The words with which the (parallel) activation of concepts is described in ACT* recalls the description of the content of working memory (Cowan, 1988, 1995, 1999; Oberauer, 2002, 2009b; Oberauer & Lange, 2009; Oberauer et al., 2013). Therefore, it is a legitimate assumption that related concepts that are active simultaneously in working memory should mutually facilitate each other's activation. This assumption can be made for several kinds of semantic similarity, including evaluative congruency.

1.1.3 Compound-cue theories and the retrieval account

As with spreading activation theories, *compound-cue theories* (Dosher & Rosedale, 1989; Ratcliff & McKoon, 1988) have been used to explain semantic priming, for example, in the lexical decision task. Generally, compound-cue theories are models about memory access and the content of retrieval cues (Ratcliff & McKoon, 1988, see also McNamara, 2005). When they are applied to priming, it is assumed that there is a memory-cue containing the target item as well as elements of its context, which can include the prime. Therefore, prime and target are in some sense active in parallel. Although this might relate to the first assumption of the three-process model by Schmitz and Wentura, that is, parallel activation of prime and target (Schmitz & Wentura, 2012; Schmitz et al., 2014), compound-cue theories do not assume a mutual facilitation. It should also be noted that it is often stated that compound-cue theories only predict beneficial effect of relatedness for material that is stronger semantically associated than being merely evaluatively congruent (Schmitz, 2012). Nevertheless, compound-cue theories address memory access and their application to working memory or, at least, their combination with theories dealing with working memory seems plausible per se. In fact, to explain priming effects, the compound-cue framework has to be combined with a memory theory (McNamara, 2005). Therefore, compound-cue theories can be considered as potential alternative explanations for effects of semantic relatedness on working memory performance. A theory that is potentially relevant when it comes to linking explanations of priming and short-term memory (or working memory) is the compound-cue theory by Ratcliff and McKoon (1988).

The compound-cue theory by Ratcliff and McKoon (1988). In their theory, Ratcliff and McKoon (1988) assume that prime and target join together to form a compound-cue. The familiarity of this compound-cue can vary, and compounds with related items lead to a higher familiarity signal than unrelated stimuli. For lexical decisions, the response that the target is a word can be based on familiarity. This response can be given faster the greater the familiarity of the compound-cue is. Therefore, related stimulus-pairs lead to faster response latencies than unrelated stimulus-pairs. This mechanism "involves no temporary activation of the long-term memory system." (Ratcliff & McKoon, 1988, p. 385). The authors assume instead that the prime and the target are combined and matched together against long-term memory. Therefore, Ratcliff and McKoon (1988) assume that short-term memory (or working memory) processes cause semantic priming. Thus, applied to working memory, compound-cue theories like the one by Ratcliff & McKoon (1988) combined with a proper memory model might provide a mechanism different to spreading activation that can predict better performance when memory for related compared to unrelated stimuli is tested. This holds especially true for tasks such as change detection, in which a response can be based on a familiarity signal. For the earlier mentioned list memory-tasks, in which mostly free recall is used (Cowles, Garnham, & Simner, 2010, Experiment 2; Goh & Goh, 2006; Murdock & Vom Saal, 1967; Oberauer, 2009b; Saint-Aubin et al., 2005), as well as the studies by Davelaar and colleagues (2006), explanations focusing on differences in the strength of the familiarity signal seem less well suited. This argument follows the critique of compound-cue theories in priming research. There, these theories have been criticized because they primarily explain priming effects that are obtained using binary decision tasks (e.g., lexical decision) in which a familiarity signal can also be used to give the right response. However, in order to apply compound-cue theories to priming tasks like naming and perceptual identification or to working memory tasks that require naming responses, compound-cue theories, especially the theory by Ratcliff and McKoon (1988), need modification or further specification (McNamara, 2005). An exception might be the model by M. S. Humphreys, Wiles and Bain (1993), which is described below.

The compound-cue theory by M. S. Humphreys and colleagues. M. S. Humphreys and colleagues (1993) provide a compound-cue theory that may be able to explain semantic priming in the naming task. Their model states that when there are two cues to generate a response (e.g., the prime and the target), each can generate an associative set. When a specific candidate lies in the *intersection* of the generated associative sets of both cues and if the amount of candidates in the intersection is quite small, this candidate will likely be reported. For example, the two cues *a mythical being* and *rhymes with post* can be joined together to generate a specific candidate that lies in the intersection of both sets that were generated by the single two cues. Thus, the word in the intersection, *ghost*, is generated with a high probability. The priming effect in the naming task is explained similarly. For associated prime-target pairs, the naming response of the target will lie in the intersection of the two sets generated by prime and target. When prime and target are unassociated, only the set generated by the target can contain the target, but not the set generated by the prime. Accordingly, for unrelated prime-target pairs, the intersection of sets would not contain the target. Therefore, only for related prime-target pairs is the naming response of the target facilitated. This model by M. S. Humphreys and colleagues (1993) can also explain evidence for a parallel activation of prime and target in priming experiments using tasks like naming or perceptual identification. Therefore, it is also a promising candidate to explain influences of semantic relatedness on working memory performance that cannot easily be explained by a higher familiarity signal, for compounds containing related instead of unrelated concepts. When two related words are remembered, both lie in the intersection of the sets generated by them. Thus, in a list memory task, the two related words are likely to be reported more often than two unrelated words. Again it is debatable whether this explanation would apply to effects of evaluative congruency or not. Further, the theory does not assume a mutual facilitation that can be explained by spreading activation processes, as suggested by Schmitz and Wentura (2012).

The retrieval account by Whittlesea and Jacoby (1990). The retrieval account by Whittlesea and Jacoby (1990), which is a theory that is related to the model by M. S. Humphreys and colleagues (1993), is also suited to explain a parallel activation of prime-target pairs (see also Bodner & Masson, 2003; Masson & Bodner, 2003, for a further development). Resembling the compound-cue theories mentioned before, the essential new point of the retrieval account is that the prime is only utilized when it is of functional value for target processing. Accordingly, in priming, especially for degraded targets (e.g., pLAnT) of related prime-target pairs (e.g., if *green* is the prime), the prime and fragments of target processing are incorporated into a compound-cue to retrieve the correct target representation. Therefore, the retrieval account predicts beneficial effects of semantic similarity in priming (and memory paradigms), especially when stimuli are presented degraded, masked, or only for a short duration. To reconstruct what was previously partially perceived, the compound of the initial processing results of prime and target (i.e., fragments of prime and target) is a potent retrieval cue for both items when they are related, but it less helpful for unrelated stimuli. For priming, the retrieval account predicts that degradation of the target-stimulus can *create* priming. It is further assumed that for related prime-target pairs with a degraded target, both stimuli form a higher order unit or – in other words – prime and target are *unitized*. This unitization, that is, the formation of a compound, creates priming. As for other compound-cue theories, it is debatable whether the retrieval account would apply to effects of evaluative congruency. Again, this theory seems incompatible with the assumption of a mutual facilitation that can be explained by spreading activation processes as suggested by Schmitz and Wentura (2012).

In addition to spreading activation theories and compound-cue theories, there are numerous other theories that are used to explain semantic and evaluative priming effects. Note that some of the other priming theories are based on the previously mentioned spreading activation theories, on compound-cue theories, or on parallel distributed processing models, which are described later. They are either versions of these models, or they incorporate features of these theories, as for example the hybrid model by Neely and Keefe (1989, for a more extensive overview on theories explaining semantic priming see McNamara, 2005). Therefore, most other theories lead to predictions for the experiments reported below that are similar to the predictions arrived at by the models described here in more detail.

Before discussing evidence for the assumption of a mutual facilitation in priming and working memory, theories from memory research that have the potential to explain effects of semantic relatedness will be described, among them the dual-store neurocomputational model by Davelaar and colleagues (2006) whose relevance was already mentioned.

1.2 Theoretical perspective on effects of similarity in memory research

1.2.1 The dual-store neurocomputational model

Interestingly, for memory performance, mutual support for semantically similar concepts was modeled in the dual-store neurocomputational model by Davelaar and colleagues (Davelaar et al., 2005, 2006; see also Haarmann & Usher, 2001; Usher & Cohen, 1999). Basically, the dual-store neurocomputational model makes explicit what is plausible to assume based on considerations about spreading activation theories. In studies of testing list-memory, the authors of the model observed beneficial effects of semantic relatedness on memory performance in both long-term and short-term memory, which fits the predictions by their computational model. This model, by Davelaar and colleagues (2006), consists of two interconnected layers: One layer represents the lexical semantic context, whereas the other layer represents the episodic context. In the lexical-semantic layer, self-excitatory connections, as well as global inhibition, are assumed. Together, these features implement a capacity-limited buffer that allows only a few items to be active concurrently. Thus, short-term maintenance is assumed to be based on activation-based processes. In contrast, long-term retention is explained by weight-based processes. Most importantly, within the lexical-semantic layer, excitatory connections between units representing semantically related items are assumed, which lead to a spread of activation to related concepts. It is assumed that when two semantically related units are concurrently active, the excitatory connections between them lead to mutual support, which partially offsets the global inhibition. Therefore, the model predicts relatedness-based performance enhancements, for example, in list memory tasks. Critically, this effect originates in a limited capacity short-term buffer, i.e. short-term memory or working memory. This is even the case when the effect arises in a long-term memory task instead of a short-term memory task. That is because the model assumes that items activated for longer durations acquire stronger episodic memory traces. Thus, there can also be beneficial effects of semantic relatedness in the long-term episodic-memory component. Nevertheless, this effect mainly originates in the short-term buffer. The model predicts that mutual facilitation (in the limited capacity short-term buffer) also leads to the effect that associated but not unrelated items tend to be displaced together. In addition, in the model, the beneficial effect of semantic relatedness continues: At retrieval, related items can "prime" each other, which can lead to a report of associated stimuli in clusters when free recall is used.

Interestingly, to our knowledge Davelaar and colleagues (2006) did not yet apply their model to classical semantic priming or evaluative priming data; it might be worthwhile to do so because there is a close similarity to the three-process model by Schmitz and Wentura (2012) and to compound-cue theories. Furthermore, the theory by Davelaar and colleagues is essentially a spreading activation theory. It can account for parallel activation but only for a limited number of concepts. Therefore, it is not a spreading activation theory that can be linked to working memory research, it is a spreading activation theory that incorporates a working memory module. Thus, the model can account for the first process by Schmitz and Wentura (2012). The second process by Schmitz and Wentura, i.e. mutual facilitation, can also be explained by the model by Davelaar and colleagues (2006). The connection to the retrieval account by Whittlesea and Jacoby (1990) is that the dual-store neurocomputational model (that is based on spreading activation processes) also accounts for a kind of chunking that occurs for related but not unrelated words: The dual-store neurocomputational model can predict that associated but not unassociated words tend to be forgotten or remembered together. Thus spreading activation processes (which are implemented in the dual-store neurocomputational model) might be capable to explain the *unitization* that was assumed by Whittlesea and Jacoby (1990) for related prime-target pairs (with degraded targets). As the statements above make clear, we understand the dual-store neurocomputational model as a spreading activation theory. Therefore, when it is stated that spreading activation theories can account for specific data, this also applies to the dual-store neurocomputational model that incorporates a spreading activation process.

1.2.2 Theories inspired by the examination of perceptual overlap

While the dual-store neurocomputational model by Davelaar and colleagues (2006) directly addresses effects of *semantic* similarity in working memory, there are other theories and considerations that address other kinds of similarity in working memory, such as perceptual overlap. Obviously, these theories do not directly deal with semantic or conceptual overlap; however, most of these conceptualizations could – in principle – also be applied to semantic overlap.

Lin and Luck (2009) used a classical working memory paradigm, the color change detection task. They observed better performance when similar, compared to dissimilar, colors had to be remembered. They provide several explanations of this effect that assume processes potentially applicable to semantic overlap. The first explanation suggested by Lin and Luck (2009) involves inhibitory interactions within color space. They propose that due to these inhibitory interactions, a *sharpening* of the memory representations in color space could take place when representations of similar colors are maintained. Because of this sharpening, the representations are assumed to be activated more precisely. Theoretically, this process could also take place in semantic space, e.g. leading to a more precise activation of the concept *Bentley* when it is processed together with a similar concept like *Porsche,* because these two car brands must be distinguished. When the word *Bentley* is presented together with the word *suitcase,* no sharpening would occur and neither concept would be activated more precisely. Perhaps in the latter case, one would only remember that the name of a car brand was presented, not knowing which one it was. A second, similar explanation that Lin and Luck (2009) provide is that one memory representation could serve as an anchor point for another representation and reduce a drift of the representations (in color space) with progressing time. In this case, a similar color could provide a better anchor than a dissimilar color, leading to better memory performance when similar colors are remembered. Again, this process could be assumed similarly in semantic space. The third suggested explanation is that, during maintenance, it is easier to attend to only a small region of color space (in trials with similar colors) compared to an allocation of attention to several color space regions (in trials with dissimilar colors). As with the processes described before, this process could also operate in semantic space. However, as far as we know these considerations had not yet been applied to semantic space and, accordingly, an empirical test of an application to semantic space is yet missing. When applying these theories to semantic space, they could account for parallel activation of more than one concept as well as for processes improving performance for semantically similar items. Therefore, these theories might also provide an explanation for effects of semantic relatedness that is different to mutual facilitation based on spreading activation processes.

Another interesting theory that can explain effects of perceptual influences on working memory performance is the rate-distortion theory (Berger, 1971; Shannon & Weaver, 1949). Sims, Jacobs, and Knill (2012) used this theory to explain why participants stored similarly oriented arrows, or lines with similar length, better than the corresponding dissimilar stimuli. They compare working memory to an information transmission channel. More precisely, they suggest that performance in a working memory task can be described as similar to sending a signal from one location to another location with distortion of the signal by noise. For memory tasks, the same principles would apply with the only difference that information is not sent across distances, but rather information has to be efficiently stored and transmitted across time, from the present to the future. Applying the rate-distortion theory to their working memory tasks, they suggest that features drawn from a distribution with a lower variance can be stored using fewer bits (pieces of information). Therefore, when there is a capacity bottleneck, the features drawn from a distribution with a low variance are tremendously easier to encode, which results in higher memory precision. Assuming that semantic space can be conceptualized similarly to the perceptual feature space, this theoretical framework could also apply to the storage of semantically similar and dissimilar concepts. As with all theories considered in this section, the rate-distortion theory does obviously allow for the parallel activation of several concepts. However, there is no direct, obvious link of the process they suggest to the mutual facilitation of related concepts, as was assumed by Schmitz and Wentura (2012).

Please note that in all of the studies conducted to test the theories mentioned here, the features defining perceptual similarity were task-relevant. The studies cannot provide insight into whether or not perceptual similarity has an effect when the feature defining perceptual similarity is not task-relevant.

1.3 Interference and inhibition

While all of the theories described above indicate the presence of automatic performance boosts by semantic similarity, there are also effects and theories indicating that drops in performance due to semantic relatedness can occur. Prominent examples of the potential presence of interference between related stimuli or inhibition of related concepts are the *picture-word interference paradigm* (e.g. Glaser & Düngelhoff, 1984; La Heij, 1988; Schriefers, Meyer, & Levelt, 1990) and the *center-surround inhibition* mechanism as proposed by Dagenbach et al. (Carr & Dagenbach, 1990; Dagenbach, Carr, & Barnhardt, 1990). In this context, semantic similarity usually leads to negative effects (e.g., slower response times). That is, in the picture-word interference paradigm, related distractor word stimuli that are presented together with a picture typically slow down picture naming response times, relative to unrelated distractor words (Glaser & Düngelhoff, 1984). Please note however, that for the picture-word interference paradigm (Glaser & Düngelhoff, 1984), it can be assumed that there is a processing-*advantage* for semantically related stimuli that enhances Stroop-like interference (e.g. due to conflicting naming responses) in trials with semantically related items. Thus, this effect does *not* provide evidence against theories suggesting mutual facilitation for semantically similar concepts. For center-surround inhibition in priming, a different argument is relevant. In studies observing center surround inhibition, and therefore negatively signed effects of conceptual similarity, the primes are not identified by the participants and therefore they are only weakly activated (e.g., Dagenbach et al., 1990; Frings et al., 2011). Accordingly, there are some hints that negatively signed priming effects are observed especially when the prime does *not* enter working memory. Therefore, the negatively signed priming effects in these studies do not provide evidence against the notion of a mutual facilitation of concepts that are concurrently in an active state in working memory.

Interestingly, there is a series of experiments focusing neither on center-surround inhibition nor on picture-word interference that observed negative effects of semantic similarity (Shivde & Anderson, 2011). The study by Shivde and Anderson (2011) deals with both the effect of semantic relatedness and the working memory system. More precisely, the authors suggest the existence of a *semantic working memory*. Therefore, a semantic maintenance capacity is assumed that is independent of visual or phonological maintenance. Participants in their studies saw a word and had to remember the meaning of the concept. Later, they had to state whether a presented target had the same meaning or not. During meaning maintenance, several stimuli were presented, with which participants had to perform a lexical decision task. One of these words was a probe that was either related or unrelated to the maintained meaning. Which of these stimuli was the probe, was however not noticeable for the participants. The authors report evidence for slower lexical decisions on semantically related probes than on semantically unrelated probes that were presented during maintenance. This can be regarded as evidence for inhibition of a concept semantically related to the maintained item. However, Shivde and Anderson (2011) assume two potential preconditions for inhibition to occur in semantic working memory. First, inhibition seems to unfold over time, like proactive interference, which tends to be smaller at shorter intervals (Hofer, 1965; Loess, 1964). Accordingly, longer maintenance intervals should be needed to observe inhibition. Second, they assume that there was a special need for inhibition introduced by the task requirements. In their paradigm, only one word had to be remembered, whereas the other words including the probe were distractors with regard to the memory task. This means that it is not inhibition of a concept maintained in working memory that was observed. Instead, participants had to inhibit the words presented during maintenance (with which they performed a lexical decision task) to *prevent* them from intruding into the memory. This inhibition could be higher for semantically related items because they were initially automatically activated by the system more strongly. Therefore, the inhibitory processes might originate in an initial performance boost due to semantic similarity that turns into an inhibitory process over time, and arguably, the inhibition occurs only under specific task demands (Shivde & Anderson, 2011). Therefore, the findings and theoretical considerations by Shivde and Anderson (2011) are also not in direct contradiction to the assumption of a mutual facilitation of semantically related concepts in working memory.

To sum up, there are theories implying both enhancing effects of semantic relations on working memory performance as well as hindering effects by inhibitory processes. In general though, as long as there is no specific need for an inhibition of semantically related concepts and no strong response conflict, or similar influences, the theories considered so far predict a performance boost due to semantic relatedness than costs or interference for semantically similar concepts. Therefore, in working memory, better performance due to similarity, for example caused by mutual facilitation, seems to be a more plausible assumption than reduced memory performance for similar concepts due to interference processes. Still, depending on the task, performance might be better when dissimilar concepts instead of similar concepts have to be remembered in a working memory task. This could be the case, given that there is a need for inhibition and time for inhibitory processes to unfold (see Shivde & Anderson, 2011) or maybe when concepts must be somehow separated, what could be difficult when (for semantically similar concepts) a higher order unit was formed (Whittlesea & Jacoby, 1990). These potential moderators of the effect of similarity can provide valuable contributions for the further investigation of the precise mechanisms that underlie effects of semantic relatedness in both priming and working memory research. Nevertheless, using “standard” procedures, working memory performance should be boosted by semantic similarity.

The theories reviewed here suggest that it is a worthwhile endeavor to merge priming and memory research. A promising starting point might be to use a task originating from priming research and to enrich this task by a memory component. In such an implementation, an enhancing effect of semantic relatedness would most likely be revealed. Studies investigating the effects of relatedness with a classical working memory task can then be used to provide evidence for similar underlying mechanisms for effects of semantic relatedness in priming and working memory paradigms. For both approaches, influences of guessing strategies and the impact of perceptual overlap should be minimized or controlled.

Next, however, the existing evidence for parallel activation in priming, and for the automatic mutual facilitation of related concepts in priming and working memory research is considered, along with an examination of the paradigms which might prove fruitful to measuring these processes.

1.4 Existing evidence and paradigms

1.4.1 Evidence for parallel activation and mutual facilitation in priming

To provide insight into the evidence for mutual facilitation of concurrently active concepts in priming that was assumed by Schmitz and Wentura (2012), we start by splitting the mutual facilitation into its two pieces. Mutual facilitation in a narrow sense means that the prime helps to encode the target and the target helps to maintain the activation of the prime. The beneficial influence of a related prime on target-encoding is also labeled proactive priming, and the reversed influence is often called retroactive priming (Dark, 1988; VanVoorhis & Dark, 1995). Observing both in the same paradigm can be considered as evidence for mutual facilitation. Further, when participants perform specific tasks on *both* stimuli, a certain performance level suggests additionally that more than one single concept is activated. This relates to the first process assumed by Schmitz and Wentura (2012) that is a parallel activation of prime and target. Thus, the following paragraphs follow this logic by first targeting automatic proactive priming and then retroactive priming, followed by a closer look at the combination of both: a mutual facilitation of simultaneously activated related concepts.

Automatic proactive priming. The semantic priming effect is a well-established and robust finding in psychological research (see McNamara, 2005 for an overview). As described previously, both definitions and procedures of semantic priming consider that the prime *precedes* the target. Also, the existence of *automatic* priming effects is often assumed. As McNamara states: “Semantic priming almost certainly is not caused solely by strategic processes" (McNamara, 2005, p. 65). Priming effects are even observed if participants do not attend to the primes (Fuentes, Carmona, Agis, & Catena, 1994; Fuentes & Tudela, 1992). Semantic priming effects are observed even when the occurrence of related prime-target pairs cannot be expected by participants (Fischler, 1977). Evidence for priming in the absence of strategic processes is typically observed especially for short stimulus onset asynchronies (SOAs; see de Groot, 1984; den Heyer, Briand, & Smith, 1985; Neely, 1977, 1991). Nonetheless, priming effects can also be strategic. Participants can, for example, generate candidates for the target that follows a related prime, which is often described as an *expectancy* based strategy (C. A. Becker, 1980). Besides this, there is also another strategic influence that can be found in the lexical decision task. This process is called *semantic matching*. In lexical decision tasks, participants have to decide whether a target is a word or a non-word. Non-words do not have a meaning; therefore, only words can be related. This correlation can be used to respond "word" when a semantic relation is detected. As a result, there can be a bias to respond "word" in related trials (Balota & Lorch, 1986; de Groot, Thomassen, & Hudson, 1982, 1986; Neely & Keefe, 1989; Seidenberg, Waters, Sanders, & Langer, 1984). This strategy is also called *postlexical* relatedness checking because the relatedness can only be detected after both stimuli have been identified. Importantly, this strategy cannot be used when a task like the naming task or the perceptual identification task (Pecher, Zeelenberg, & Raaijmakers, 2002) is utilized. Accordingly, these tasks are suitable candidates to provide evidence for automatic proactive priming and possibly also for automatic retroactive priming and automatic mutual facilitation of related concepts using priming paradigms.

While the presence of semantic priming effects using naming, or other tasks different than lexical decision, is nearly unquestioned, and also the existence of *automatic* semantic priming effects is broadly accepted, the existence of automatic S-S-based *evaluative* priming effects has recently been questioned. While Bargh, Chaiken, Raymond and Hymes (1996) observed S-S-based evaluative priming effects using a naming task (see also Hermans, Houwer, & Eelen, 1994, Experiment 2), this effect was not replicated in several carefully conducted replication attempts by Klauer and Musch (2001), as well as by Spruyt, Hermans, Pandelaere, De Houwer and Eelen (2004). The ambiguity as to whether or not there is automatic S-S-based evaluative priming could perhaps be explained by differences in the allocation of attention by participants across tasks. Naturally, in the evaluative decision task participants attend to the evaluative component of the stimuli. In contrast, when a naming response is required, there is no need to focus on the valence of the presented stimuli. This central difference between the evaluative decision task was addressed by the feature-specific attention allocation framework by Spruyt and colleagues (Spruyt, De Houwer, & Hermans, 2009; Spruyt, De Houwer, Hermans, & Eelen, 2007). Their theory predicts that it is mandatory in evaluative priming to assign attention to the affective stimulus dimension in order to observe a significant evaluative priming effect. In a priming study, they observed that when participants had to perform an evaluative decision task in the majority of trials, an evaluative priming effect emerged in trials with the evaluative categorization task as well as in trials using a non-affective semantic categorization task. On the other hand, the evaluative priming effect disappeared (even in the evaluative decision task) when participants assigned attention to non-affective stimulus features in 75% of trials (Spruyt et al., 2007). Please note that while Spruyt and colleagues regard the allocation of attention to the valence of the stimuli as mandatory, some early studies observed evaluative priming effects irrespective of a focus on valence (e.g. Bargh et al., 1996; Hermans et al., 1994). Furthermore, Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008), observed significant S-S based evaluative priming in several experiments. Interestingly, the latter studies, which will be described in a later section, include experiments with a focus on parallel activation and mutual facilitation of evaluatively congruent prime-target pairs. Further, there is also meta-analytical support for evaluative priming effects that are not obtained using the evaluative decision task and therefore measured in tasks in which usually no strong focus on valence is given (Herring et al., 2013). Nevertheless, the degree to which the valence of stimuli is processed might be a valuable moderator of the evaluative priming effect, which might explain why picture stimuli usually elicit stronger evaluative priming effects (see Herring et al., 2013).

Retroactive Priming. An often neglected aspect of semantic and evaluative priming is the influence that the processing of a related target concept can have on the processing of the prime concept. With this in mind, the following paragraph does not focus on effects that can be explained by a facilitating influence on target identification. Instead, studies are considered that analyzed the reversed effect: the effect of related targets on the prime. Such effects are often also labeled *retroactive priming* (Briand et al., 1988; Dark, 1988), a term that we use here as a general label for all backward directed priming effects when the target has an influence on prime processing.

An interesting finding in the literature is that after the usual response to the target, the primes are reported more often in related compared to unrelated trials when the prime has to be named (Dark, 1988; VanVoorhis & Dark, 1995). Because in these experiments both is observed, retroactive and proactive priming, they also provide evidence for *mutual* facilitation. However, for these studies, there is an alternative explanation for the observed effect. The retroactive priming task in which a subsequent prime report is requested could be regarded as two overlapping proactive priming tasks. First, a related prime leads to facilitated target encoding. Second, the (old) target serves as a prime for the original prime that has to be retrieved and is now the new target. Therefore, this task would not unequivocally show retroactive priming as a backward directed process (VanVoorhis & Dark, 1995). Thus, it is not necessarily the case that the analysis of the target maintains the activation of a related prime. Assuming for a moment that the evidence for mutual facilitation (of concurrently active related concepts) is substantial and the potential alternative explanation does not solely explain the data, the task has an interesting aspect: By instructing the participants to recall the prime, a (working) memory component is incorporated into the design of this priming task. Another legitimate description of the retroactive priming effect is to assume better memory for word stimuli that were followed by a related word. Please note that in the retroactive priming studies mentioned above, the prime can in most trials be named after a response to the target was given. Further, in studies not using the lexical decision task to provide a response on the target, the target also has to be named (Dark, 1988; VanVoorhis & Dark, 1995). Therefore, when the prime and the target can both be identified in most trials, both concepts are active concurrently. Accordingly, in principle, the experiments on retroactive priming provide evidence for the first two processes of the three process model by Schmitz and Wentura (Schmitz & Wentura, 2012; Schmitz et al., 2014): First, prime and target can be active simultaneously, and second, there is a mutual facilitation of related prime-target pairs.

A design similar to the retroactive priming paradigm was used by Whittlesea and Jacoby (1990) to test their retrieval account described above. They presented first a prime and afterward a target (which they termed interpolated word). After that, there was a second target (transfer target), which matched the prime (in trials in which primes were presented instead of a neutral stimulus – "@@@@@"). Their theory predicted faster response latencies for the second target (the repetition of the prime) when the first target (the interpolated word) and the prime are related and when the first target is additionally degraded. This finding was interpreted as indicating that the prime is more included in the analysis when prime and the (first) target are related and when the (first) target is degraded. In this case, the related prime can be used to identify the difficult to encode degraded target. Accordingly, their experiments show also a "retroactive", backward directed priming effect that is, however, dependent on the degradation of the (first) target. Assuming that the prime is more included in the analysis, the authors conclude that related prime-target pairs with a degraded target are *unitized*. Because of the moderation of this retroactive priming effect by the degradation of the target, the effect can convincingly be explained by the retrieval account and the assumption of a higher unitization as the cause of the enhanced performance.

Besides the effect found in the study by Whittlesea and Jacoby, retroactive priming effects can easily be explained by mutual facilitation based on a spread of activation. Also some compound-cue theories as well as theories inspired by assumptions about the effect of perceptual overlap can account for unmoderated retroactive priming. These explanations as well as explanations by spreading activation theories might need additional assumptions to explain the effect of target degradation on retroactive priming in the study by Whittlesea and Jacoby (1990). As already mentioned, parallel distributed processing models, such as the model by Masson (1991, 1995), rely on the temporal order of prime and target activation and might, therefore, need to be modified to explain backward directed effects (see Schmitz & Wentura, 2012; Schmitz et al., 2014 for this argument).

Mutual facilitation. As outlined above, some retroactive priming studies apparently already provide evidence for a mutual facilitation of simultaneously active related concepts. The most convincing evidence for a mutual facilitation of semantically related concepts, however, stems from the research by Wentura and colleagues on evaluative priming. Wentura and Frings (2008), as well as Schmitz and Wentura (2012, see also Schmitz et al., 2014), tested semantic priming effects by manipulating the overlap of the evaluative component of stimuli (that is, prime and target were either different in their valence or both positive/negative). In these studies, negative SOAs were used to foster the influence from the evaluatively congruent targets on the maintenance of the prime and to minimize the effect of evaluative congruence on target encoding. A positive priming effect (faster response latencies in the case of evaluative congruency) was only observed when prime pictures were not closely associated with naming responses (Schmitz & Wentura, 2012, Experiment 1; Wentura & Frings, 2008). On the other hand, priming was either absent or significantly reversed when primes were strongly associated with naming responses, which could interfere with the response to the target. This finding can be explained by mutual facilitation of prime and target. First: a related prime helps the target to be encoded (proactive priming). Second, a related target helps to maintain the prime activation (retroactive priming). Usually, the second process is of no relevance; however, when the prime triggers a naming response that interferes with the naming of the target, there is interference of both responses, slowing the correct response down. In the case of a related (evaluatively congruent) prime, the activation of the prime is maintained more than for unrelated (evaluatively incongruent) primes. Therefore, there is more response conflict by the activation of the interfering response to the prime in a congruent trial, resulting in the reversed priming effect when response bound primes are used. Similar interference effects due to mutual facilitation of related, simultaneously active concepts were obtained in Experiment 2 & 3 by Schmitz & Wentura (2012), as well as with the flanker paradigm by Schmitz, Wentura & Brinkman (2014). There, a semantic categorization task was used. In Experiment 2 by

Schmitz and Wentura (2012) for example, primes and targets were evaluatively positive and negative pictures. Further, these pictures showed persons or animals. Because a categorization response was required for the non-evaluative dimension (i.e., person vs. animal), it was possible to vary evaluative congruency orthogonally. Thereby, there are evaluatively congruent trials as well as evaluatively incongruent trials with and without response conflict, as long as it is assumed that prime and target (or at least the responses associated with both stimuli) are active simultaneously. In this experiment, again, a negative SOA was used. The authors observed a significant interaction of the factor evaluative congruency and the factor indicating the overlap in the non-evaluative dimension, which also indicates the presence or absence of response conflict. In other words, only when prime and target matched with regard to their valence was there a significant S-R-based semantic priming effect. The effect was however not observed when prime and target did not match regarding their valence, because the prime was too weakly activated to trigger the corresponding response. The effect can also be described in another way: In evaluative congruent trials, the parallel activation of prime and target led to a (numerically) stronger effect of response conflict than in incongruent trials when the non-evaluative dimension of the stimuli required diverging responses. When the non-evaluative dimension of the stimuli required the same response for the prime and the target, the usual pattern of a positive evaluative priming effect was observed. Importantly, there was a learning phase that ensured that the primes were associated with responses to enable response conflict. The pattern of results can be explained by the three processes assumed by Schmitz and Wentura (2012): (1) prime and target are active in parallel, and (2) they mutually facilitate each other's activation in the case of evaluative congruency. This mutual facilitation leads to fast responses when there is no response conflict. Otherwise, however, the further process (3) comes into play: the response conflict, which is especially pronounced in the case of evaluative congruency when prime and target mutually facilitate each other's activation which also boosts the associated responses and therefore the response conflict. The same pattern of results was also obtained in Experiment 3 in which words instead of pictures were used (Schmitz & Wentura, 2012) and in the study by Schmitz and colleagues (2014) in which a flanker task with words was implemented. In the latter study, EEG was also recorded and an

ERP component called the N2 was measured, the amplitude of which can be regarded as a marker of response conflict (van Veen & Carter, 2002). Thereby, evidence was gathered that the pattern of results was indeed due to enhanced response conflict due to the maintenance of prime activation by evaluative congruency.

Interestingly, the effect observed with this specific paradigm and these specific stimuli used by Schmitz and Wentura (2012) occurred only when relatedness was induced via the overlap of the *evaluative* component (and when the person-animal distinction was task-relevant) and *not* when evaluative congruency was task-relevant. That is, in their task, similarity on the person-animal dimension did *not* cause mutual facilitation (see Experiment 4a/b). This may indicate that mutual facilitation of related concepts is a valence-specific mechanism. Alternatively, this finding could be caused by a higher salience of the evaluative component in the materials. That is, in the experiments reported by Schmitz and Wentura (2012), stimuli were selected due to high valence ratings (either in a pilot study by De Houwer, Hermans, Rothermund, & Wentura, 2002 or because of high valence ratings in the International Affective Picture System, Center for the Study of Emotion and Attention, 1994). Thereby, the material itself might have introduced a focus on the evaluative dimension, which according to Spruyt and colleagues (2009, 2007) is required to obtain evaluative priming effects. The distinction between animals and persons might not have such relevance especially when stimuli are selected because of high valence ratings. Furthermore, taking the evidence from other studies into account it seems rather unlikely that mutual facilitation is restricted to stimuli sharing their valence.

It should be mentioned again that, in the tasks used by Schmitz and Wentura (2012), when there is a response conflict between prime and target, both stimuli must be active at the same time, in parallel. At least this is the case for the responses associated with both stimuli. Therefore, they provide evidence for both: a parallel activation of prime and target and a mutual facilitation in the case of evaluative congruency. When there is evidence that prime and target are concurrently active, a link to working memory research suggests itself, because it is typically assumed that up to four items can be concurrently active in working memory (e.g. Alvarez & Cavanagh, 2004; Cowan, 2001; Luck & Vogel, 1997).

The evidence by Schmitz and Wentura (2012) can parsimoniously be explained by the three-process model they suggest. On a process level, the most suitable explanation of their findings might be the theories assuming a spread of activation (J. R. Anderson, 1976, 1983, 1993; Collins & Loftus, 1975), which also include the dual-store model by Davelaar and colleagues (2006) originating from memory research. The compound-cue models do not need the assumption of mutual facilitation to explain parallel activation of prime and target as well as beneficial effects of semantic overlap in priming. To apply these theories to the effects obtained by Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008), the compound-cue model by Ratcliff and McKoon (1988) that focuses on an enhanced familiarity due to similarity cannot directly explain the findings of the experiments in which naming responses were required. For the compound-cue model by M. S. Humphreys and colleagues (1993), and the retrieval account by Whittlesea and Jacoby (1990), it remains unclear whether these theories can explain the evidence for parallel activation and mutual facilitation obtained by Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008). To explain these findings, it must still be determined how the formation of a compound would introduce responses associated with the prime and the target. To our knowledge, no precise predictions about the combined effects of negative SOA, response conflict and the role of evaluative congruency (and semantic similarity) can yet be derived from these models. Further, Schmitz (2012) states that the existence of S-S-based evaluative priming generally might be difficult to explain for compound-cue theories, because compound-cue theories usually rely on a direct association and a frequent co-occurrence of the prime and the target concept. Evaluatively congruent concepts are however not per se associated in that way.

Bridging the gap: Working memory components in priming paradigms. As pointed out earlier, the retroactive priming task by Dark and colleagues (Dark, 1988; VanVoorhis & Dark, 1995) has a worked-in memory component. The prime has to be reported and therefore remembered after the target was presented and semantically related targets seem to help maintain the prime. Therefore, and because participants are also able to identify the target, their results provide a hint for parallel activation of prime and target and mutual facilitation in the case of relatedness. Furthermore, as derived by Schmitz and Wentura (2012) for their own findings, the results by Dark and colleagues can also be interpreted as a sign that working memory models, which can account for the parallel activation, could be extended to additionally account for mutual facilitation of related concepts.

Obviously, the experiments by Schmitz and Wentura (2012; see also Schmitz et al., 2014) do not directly incorporate a (working) memory component in the sense that something has to be remembered. Nevertheless, the evidence for the parallel activation of prime and target does also suggest a link between priming and working memory research.

Similarly to the retroactive priming studies, another paradigm can be used to incorporate a working memory component into priming paradigms: the post-cue task. In this task pairs of stimuli (for example words) are presented. Therefore, the paradigm resembles a semantic priming task with SOA zero. However, in contrast to classical semantic priming studies, the to-be-reported stimulus is not defined before the offset of the stimuli (see, e.g., Dallas & Merikle, 1976a, 1976b; G. W. Humphreys, Lloyd-Jones, & Fias, 1995; Murphy, 2010; Murphy & Green, 2011). A common finding is better performance, e.g., faster naming latency, when the two stimuli are related (but see G. W. Humphreys et al., 1995). There are three interesting implications of the post-cue task. First, using the post-cue task the usual sequence of a prime presented prior to a target is overcome by using a 0 ms SOA. Second, there is arguably no asymmetry between prime and target at encoding; both stimuli are equally likely to be marked as the target by the post-cue. Third, if the overall accuracy in post-cue tasks using naming (or similar tasks) is above 50 %, then more than a single stimulus is active at one point in time, providing evidence for the first process assumed by Schmitz and Wentura (2012), which is parallel activation. Therefore, in this case, better performance in related compared to unrelated trials can parsimoniously be explained by mutual facilitation of related concepts. However, not all post-cue studies necessarily provide evidence for an *automatic* mutual facilitation of concurrently active related concepts. This is because most of the experiments using post-cue tasks used a late onset of the cue or rather long presentation times for the word-pair. This might have led to expectancy-based effects that are usually observed for longer SOAs in sequential priming tasks (see C. A. Becker, 1980).[3] These ideas and considerations are further discussed in the description of Experiment 1b in which a post-cue task was used.

Despite the fact that the retroactive priming paradigm might not unambiguously measure retroactive, backward directed priming effects and the results of post-cue priming studies might partially be due to expectancy based processes, studies using both tasks provide preliminary evidence for the three process model by Schmitz and Wentura (2012). These studies can also be regarded as evidence for the applicability of the dual-store neurocomputational model by Davelaar and colleagues (2006) to priming paradigms, and obviously, these effects can be explained by spreading activation theories in general. Alternatively, retroactive priming effects and effects in post-cue tasks can be explained by some compound-cue models that predict priming effects using the naming task (M. S. Humphreys et al., 1993; Whittlesea & Jacoby, 1990) and that do not assume a mutual facilitation of related concepts.

Another series of priming studies that provides a hint that working memory processes might be involved in the formation of beneficial effects of semantic similarity is the research by Mahr and Wentura (2014). They investigated the facilitating effect of time compressed auditory stimuli on the identification of the corresponding color in a visual search task, in which not only relatedness, but also perceptual load, as defined by Lavie (1995), was manipulated. Mahr and Wentura (2014) observed that related auditory primes (compared to neutral or unrelated auditory primes) facilitated the processing of a corresponding target-color patch. In a similar study by Chen and Spence (2011, Experiments 3b and 4a), however, no cross-modal priming by spoken words was observed. Mahr and Wentura (2014) state that one central difference between the studies is that Chen and Spence (2011) used pictures of *30* items while, in the study by Mahr and Wentura (2014), *four* relevant target colors were used. Mahr and Wentura (2014) conclude that the participants in Experiment 3b and 4a by Chen and Sprence (2011)

[3] Dallas and Merikle (1976b) used a short post-cue delay as well as short presentation durations; however, rather uncommon methods were used: targets were repeated six times in the experimental trials, while primes were presented only twice. Furthermore, stimulus lists were not counterbalanced.

most likely did not keep the large number of items permanently in mind. Presumably, in the study by Mahr and Wentura (2014), participants always kept the four relevant target colors in mind. It seems that a strong mental set is required to obtain a priming effect in this task. In other words, the "four relevant target items had to be actively kept in working memory throughout the experiment" (Mahr & Wentura, 2014, p. 588). Priming one color might have put this color into the "foreground" (Mahr & Wentura, 2014). The existence of such a prioritization in working memory was assumed by Olivers, Meijer, and Theeuwes (2006). The considerations by Mahr and Wentura, are thereby closely related to the explanation of the *zigzag* effect for related word pairs in immediate recall which was observed and explained by Davelaar and colleagues (2006), described in more detail in section 1.4.2. This zigzag effect originating in working memory for lists with related words is also considered to be an effect that is caused because subsequently presented stimuli enhance the activation of stimuli already maintained in working memory (e.g. Davelaar et al., 2006). In a similar manner, the considerations by Mahr and Wentura (2014) also relate to the assumption of a retroactive priming (e.g. Dark, 1988) and the assumption of a mutual facilitation of related prime-target pairs by Schmitz and Wentura (2012) described in previous paragraphs. These processes or effects have in common that they can be explained by most spreading activation theories (J. R. Anderson, 1976, 1983, 1993; Davelaar et al., 2006) and potentially also by some compound-cue theories (M. S. Humphreys et al., 1993; Whittlesea & Jacoby, 1990). To make this assumption for the study by Mahr and Wentura (2014), it is important to mention that they also observed a cross-modal priming effect when a target-present or target-absent decision was required. Thereby, it can be ruled out that response-priming processes caused their congruency effect. Instead, a stimulus-stimulus compatibility and conceptual overlap caused their effect, which can, therefore, be explained by spreading activation theories. Essentially, the difference in results between the studies by Mahr and Wentura (2014) and Chen and Spence (2011) suggests that an involvement of working memory processes is needed to obtain the priming effect that Mahr and Wentura (2014) observed.

1.4.2 Parallel activation and mutual facilitation in working memory?

Parallel activation: the working memory. While there is still debate about the precise number of items that can be stored in working memory, it is almost certain that more than a single concept can be maintained in working memory (but see Olsson & Poom, 2005). In a highly cited, groundbreaking paper, Miller (1956) states that seven (plus/minus 2) objects can be maintained in working memory (or short-term memory). However, in the more recent theory by Cowan (2001) a working memory capacity of about four concepts is suggested. In line with this assumption, many studies observed a capacity, e.g. in visual working memory, of around three or four chunks of information as the classical study by Luck and Vogel (1997). Notice, however, that the capacity of the working memory seems to depend on the discriminability between the stimuli (Alvarez & Cavanagh, 2004). Nonetheless, it can be stated almost certainly, that more than a single concept can be maintained in working memory. Accordingly, a contribution of working memory can most easily account for a parallel activation of several concepts in priming, an assumption that seems to be needed (together with other assumptions) to explain a number of tricky findings in priming research (see Schmitz & Wentura, 2012).

Furthermore, as already mentioned, spreading activation theories are compatible with working memory models in which the content of working memory is regarded as part of the activated concepts in long-term memory (Cowan, 1988, 1995, 1999; see also Oberauer, 2002, 2006, 2009a; Oberauer & Lange, 2009; Oberauer et al., 2013). While it is nearly unquestioned that several concepts can be maintained simultaneously in working memory, mutual facilitation of related concepts in working memory is, however, a rather uninvestigated phenomenon.

Mutual facilitation or mutual inhibition in working memory tasks? First of all, it should be mentioned that the priming studies discussed above can also be considered as evidence for mutual facilitation in working memory. However, the following paragraphs only focus on studies that used classical memory paradigms, but not, for example, priming tasks that include a memory component, although this distinction might be somewhat artificial. Nevertheless, here, studies are considered that provide either evidence for beneficial or detrimental effects of (semantic) similarity on memory performance.

Preliminary evidence from studies on other types of similarity

Firstly, one study analyzing the effect of perceptual overlap in working memory should be considered. While it may seem tempting to assume that studies from this area are mainly important to provide models that can also be applied to effects of semantic relatedness (see section 1.2.2), they also have another interesting aspect. In some studies investigating perceptual similarity, *conceptual* similarity might also have contributed to the obtained effect. Before taking a closer look at some of these studies, please realize that, in general, most studies investigating the effect of perceptual similarity, in which all stimuli stem from the same category and there is no strong semantic or perceptual dissimilarity, provide evidence for beneficial effects of similarity in working memory tasks (Jiang, Lee, et al., 2016; Lin & Luck, 2009; Sims et al., 2012). In general, at least, it can be assumed that in most of these studies there is also a slight semantic similarity that correlates with perceptual similarity, which is manipulated in these studies. That is when different shades of red are presented, these shades have in common that they all activate the *semantic* concept “red”. When shades of red, green, yellow and blue have to be remembered, participants might try to remember four distinct *concepts* instead of one. Thereby, the effect obtained by Lin and Luck (2009) could also be partially based on semantic similarity.

The same considerations apply even more to a study by Jiang, Lee and colleagues (2016), which used perceptually similar faces. They found that similarity led to a better working memory performance. In this study, perceptually similar faces were generated using morphs containing the same celebrity. Dissimilar faces were faces each containing the face of different celebrities. However, faces of celebrities are not unfamiliar faces. These faces are stimuli for which there is a vast semantic knowledge base. Therefore, the beneficial effect of similarity in their study might be caused because the same semantic concept was activated for all memorized stimuli in trials with high "perceptual" overlap.

For the study by Sims and colleagues (2012) that also found beneficial effects of perceptual similarity, there is no obvious relation to semantic similarity (e.g. for arrows with similar orientations or lines of similar length, there is no clear link to semantically similar concepts). However, the authors admit that there are several alternative explanations for their results different than an effect of perceptual similarity. They provide convincing arguments and analyses to demonstrate that none of these alternative explanations alone can account for their findings. Nevertheless, they do not rule out the possibility that several alternative explanations *combined* can fully account for their effect. Accordingly, the seemingly most convincing evidence for effects of *perceptual* similarity on working memory performance could indeed also be considered as evidence for beneficial effects of conceptual or *semantic* similarity.

Nevertheless, there is also another tradition of working memory research in which mainly disadvantageous effects of similarity were observed. This is the research on phonological similarity (e.g. Baddeley, 1966; Gupta, Lipinski, & Aktunc, 2005). Please note, however, that phonological similarity is not correlated to semantic similarity, in contrast to the visual similarity in the studies mentioned before. Therefore, there is no convincing reason to assume that the pattern of results that these studies generated will translate to the effect of semantic similarity on working memory performance.

Evidence from studies manipulating semantic relatedness and the use of common category retrieval cues

Obviously, there are other important experiments that directly target effects of semantic relatedness on working memory performance. There is a long tradition of research analyzing the effect of relatedness using semantically related and unrelated words that participants must store and remember in short-term memory tasks. In most of these studies, participants performed better when the to-be-remembered words were related (Cowles et al., 2010, Experiment 2 Goh & Goh, 2006; Murdock & Vom Saal, 1967; Oberauer, 2009b; Poirier & Saint-Aubin, 1995; Saint-Aubin et al., 2005). In other experiments, no significant effects or mixed results were observed (Baddeley & Levy, 1971; Cowles et al., 2010, Experiment 2; van der Lely & Howard, 1993), and in still other studies a reversed effect was observed (Baddeley, 1966; Dale & Gregory, 1966). Regarding this topic, the studies investigating the semantic blocking effect are also of interest. Presenting words blocked by category rather than randomly intermixed also leads to better performance (see Calfee & Peterson, 1968; Huttenlocher & Newcombe, 1976)[4]. Also for the effects of semantic blocking, it could be speculated that they originate in working memory, particularly considering that the effects were observed in working memory tasks.

Nevertheless, there is a relevant point concerning the interpretation of the beneficial effect of semantic relatedness in paradigms like list memory tasks, including the semantic blocking studies: Better performance for associated items, related items or items blocked by category can be explained by the use of categorical retrieval cues (Calfee & Peterson, 1968; Poirier & Saint-Aubin, 1995). Accordingly, the studies reported here do not necessarily measure *automatic* effects of semantic relatedness on working memory performance. Therefore, the studies of this thesis were designed in a way to avoid the use of common category retrieval cues or to control for the influences of guessing strategies.

Besides the evidence form list memory paradigms that mainly indicate better performance for semantically related items (which is however not necessarily based on an automatic effect), there are, however, also some studies that observed a reversed effect of semantic relatedness in memory tasks. One finding is that performance, not for the first, but for the second or third item in a sequence of related items, is impaired (Loess, 1967, 1968; Turvey, Cremins, & Lombardo, 1969; Wickens & Clark, 1968 see also the review by Wickens, 1970). This finding is usually interpreted as proactive interference by semantic relatedness which impairs retention. As outlined by Shivde and Anderson (2011), these studies did not measure working memory performance, because the memory test in these paradigms was after a delay in which participants performed a

---

[4] see Warrington, Kinsbourne and James (1966) for a similar finding with significantly lower performance for letters and digits than for either only digits or only letters.

distractor activity such as counting back. Furthermore, these proactive interference effects seem to disappear when testing occurs immediately (Wickens, Moody, & Dow, 1981). Interestingly, the effects interpreted by these studies as proactive interference have a close resemblance to the *zigzag* effect observed by Davelaar and colleagues (2006). In their model and for their data however, it becomes evident that there is a *benefit* for the first item of a pair of related stimuli but not costs for the second item.

Still, other studies compare the performance in a condition in which visual stimuli from semantically *and* perceptually completely different categories are used and compared with a condition in which stimuli from the same category are used (M. A. Cohen et al., 2014; Jiang, Remington, et al., 2016). For example, the memory for faces and scenes was observed to be better than the memory for only faces or only scenes. This allows, however, no strong conclusions about the effects of semantic relatedness. The effect in these tasks might be caused by the strong perceptual dissimilarity. Therefore, with regard to our purpose, studies from this tradition remain rather uninformative.

To summarize, the studies mentioned in this section partially provide hints for beneficial (as well as some disadvantageous) effects of similarity. However, a direct transfer to *automatic* effects of *semantic* similarity remains questionable. Nevertheless, the effects of some studies using for example semantically related words might partially measure automatic effects. Furthermore, the studies reported here also provide insight about what should be avoided or controlled when semantic similarity is investigated, that is, the use of a common category retrieval cue and strong perceptual overlap.

The effect of a maintained concept on the processing of semantically related items

In the following sections, experiments are reported, in which it seems unlikely that the use of a common category as a retrieval cue and other strategic influences or effects of perceptual similarity did influence results. A series of studies falling into this category was conducted by Shivde and Anderson (2011). They report experiments in which an effect of semantic relatedness on reaction times was observed in a working memory paradigm. In their delayed judgment task, participants had to sustain a single item that was presented in red at the beginning of a trial. When the maintenance interval ended, a target item written in blue appeared on the screen. In the critical conditions/experiments, participants had to indicate whether the maintained item and the target item shared meaning; they were explicitly encouraged to maintain the semantic *meaning* of the item. In other words, a change detection task with a focus on the meaning of the concepts was implemented. During maintenance, and also after the response to the blue target-word, participants performed a seemingly unrelated task: A stream of words and nonwords written in black was presented, and participants performed a lexical decision task for every single word. One of these words was a probe item that was either semantically related or unrelated to the maintained item. To prevent misunderstandings, for the participants it was not made clear that this probe item is of specific interest. The critical finding is that the *word* responses to the probe item were slower when it was semantically related to the maintained item compared to when it was unrelated. This effect occurred only when the probe was presented during maintenance but not after the judgment of the target. In addition, the effect emerged only when the semantic task, i.e. a comparison of the meaning of the maintained item with the meaning of the target, was used; the effect was not observed when participants performed a corresponding phonological task. The authors conclude that a specific semantic maintenance capacity exists. An interesting point is that the effect occurred only when the *semantically* similar concept was held *actively* in working memory, and it was not observed for other types of maintenance or if participants no longer tried to remember the concept. Therefore, it seems to be a highly specific working memory process induced by semantic relatedness. However, the investigated effect is not introduced by the semantic similarity of *concurrently* active concepts. In contrast, the effect affects the encoding of an item which should *not* be maintained. Once this concept enters working memory it might be automatically maintained, but this was not investigated in their study. Therefore, the slower reaction times due to semantic relatedness do not challenge the assumption of a mutual facilitation. Please note also that the task implemented by Shivde and Anderson (2011) has specific demands on inhibition: all words within the word stream have to be inhibited in their meaning because they should not be maintained. Instead, participants had the task to compare meaning between the to-be-remembered word and the target. Accordingly, the probe, which is a stimulus from the word stream, also has to be prevented from intruding into memory. Thus, the effect might depend on this specific task that requires inhibition. In Experiment 3 of their study, Shivde and Anderson (2011) further explored the time curse of inhibitory control. They assumed that the inhibition demands should grow over time when more distracting words are presented. Interestingly, they assume that there could be an initial wave of activation leading to faster response latencies on associated probes. This activation could be followed by an inhibition that grows over time. In line with this assumption they did not find a negatively signed effect of semantic overlap of the probe when it was presented earlier during maintenance; however, they also did not observe a positive effect that would reveal more compelling evidence for an initial activation. Nevertheless, an initial activation is plausible to assume, and it can be speculated that presenting the probe even earlier during maintenance in the task used by Shivde and Anderson (2011) could have revealed faster reaction times for related probes. If their experiments measured task-specific inhibition caused by an initial automatic activation, their findings are in line with one part of the mutual facilitation process that we assume (at least for concepts entering working memory). Another aspect that could contribute to the reversed effect in the task is that the probe must be more distinguished from the target than the other distractor words because the probe, but not any other distractor word, is related to the target. One might argue that the distinct color cue is enough to discriminate between probe and target. Nevertheless, the relatedness of the probe might introduce an initial tendency for a different response, which is the response to the target, that is, an old-new response regarding the match of meaning between the maintained item and the target item. Furthermore, even if the probe would not trigger a different response, the probe could have a higher salience in this context, being the only word that is related to the maintained item. If one of these processes could explain the finding by Shivde and Anderson (2011), the slower response times in the case of semantic relatedness would be highly task-specific. Further, these processes could counteract a speeding up of responses in the condition in which the probe was presented earlier during the trial. When an initial facilitation of semantically related concepts contributes to the findings by Shivde and Anderson (2011), the effect can more plausibly be explained by spreading activation theories than by compound-cue theories. This is because compound-cue theories would arguably not assume any effects (or strongly reduced effects) of semantic relatedness in this task due to the high number of intervening items (Ratcliff & McKoon, 1988). With the assumption of different weights of the items to enter the compound-cue, this conclusion is attenuated. However, an effect of the observed size and its temporal course seem rather unexpected following the ideas of most compound-cue theories. When the specific requirements of the task used by Shivde and Anderson (2011) are considered, their effects are at least not incompatible with spreading activation theories and, further, spreading activation processes can explain the suggested initial activation.

Experimental evidence for the dual-store neurocomputational model by Davelaar and colleagues

More insight on the behalf of our purposes is provided by a series of studies used for validating the dual-store neurocomputational model by Davelaar and colleagues (2006). In their task, participants learned 12 words. Trials in the related condition contained a sequence of 6 pairs of semantically associated words. In each trial of the unrelated condition, a sequence of 6 unrelated word pairs was presented. As in semantic priming paradigms, the word pairs in the unrelated condition were created by regrouping semantically associated pairs. Among other types of recall, Davelaar and colleagues (2006) used an immediate recall. They obtained evidence for mutual facilitation by related words in working memory of a two-fold nature. First, they observed better (working memory) performance in the related compared to the unrelated condition. Second, they observed a higher *zigzag* effect in the related than in the unrelated condition. That is, in the related condition compared to the unrelated condition, the first item of a pair had a somewhat higher performance compared to the second item across serial positions in the list.[5] It was observed that related stimuli support each other, and they tend to be displaced together. Thereby related word-pairs compared to unrelated

[5] For the calculation of the zigzag effect, both the performance of the first and the performance of the second item were calculated relative to the expected performance for the specific serial position in which they were presented. Then the difference in performance was calculated.

words act like they are chunked together concerning storage and displacement from working memory. In their task, the second word in a related pair receives facilitatory activation from the first word like in proactive semantic priming. In addition, the second word facilitates the activation of the first word analogous to retroactive priming effects. Therefore, for the first word, there is even a maintaining influence or a reactivation event after its offset, so that the first item of the word pair shows better recall than it would be expected on the basis of its serial position. Accordingly, this mechanism can explain the zigzag effect in the related condition. Importantly, in the model they describe and that fits the data, both effects, i.e. the overall enhanced performance in congruent compared to incongruent trials as well as the zigzag effect, originate from a limited capacity short-term buffer, the short-term memory or working memory. Please note that there is also a zigzag effect in long-term memory that has its origin however in short-term memory. In the model by Davelaar and colleagues (2006), items that are active for a longer time have stronger episodic traces. This means, the first item is maintained in short-term memory for a longer time and thereby it will have better performance in tasks like delayed free recall, which causes a zigzag effect there. When items are however separated during encoding by a distractor activity, the short-term memory influences are prevented. Therefore, there is also no long-term memory contribution to a zigzag effect when distractor activities are used because the zigzag effect in this task cannot build on the effect from the short-term memory system. Accordingly, Davelaar and colleagues (2006) did not observe a zigzag effect when using continuous-distractor free recall. Thus, the zigzag effect they observe seems to be based on what we would call mutual facilitation in working memory.

Other evidence for beneficial effects of semantic relatedness that can arise in short-term memory or working memory were obtained by Haarmann and Usher (2001). In a similar memory task, they manipulated the separation of semantically related words. Using immediate as well as delayed recall they calculated estimations of short-term memory contribution and long-term memory contribution on task performance. Thereby, they observed that, in contrast to more distant related words, related words that are adjacent to one another are more likely to remain coactive in the short-term memory module of their model. They also conclude that related items can support each other via excitatory links. This process is more likely to occur when the related stimuli are presented adjacent to each other because of the limited capacity of the short-term memory system. Thus, Haarmann and Usher (2001) have already provided evidence for mutual facilitation of semantically related items that arises in a limited capacity short-term buffer or the working memory.

Importantly, Davelaar and colleagues (2006) argue that their effects are not based on strategic influences. They used the same stimuli as Haarmann and Usher (2001): The word pairs were taken from word norms; however, each second word of a word-pair was only the response to the first one by less than 3 % of the participants in a corresponding standardization study (in which participants were required to give associated words). However, inspection of the word-pairs given as examples reveals that parts of them could be memorized consciously and strategically as a single chunk. For example, to remember the words *heavy* and *stone* participants might simply just store the concept of a *heavy stone*. *Light* and *candle* could become a *candlelight*, *spider* and *snake* might form a common category like *poisonous animals*. In the unrelated condition, none of these strategies is possible. Although strategic influences are less likely than in studies using direct associates, they cannot be completely ruled out here either. If not caused by strategic processes, their effects can most plausibly be explained by spreading activation theories and, obviously, the dual-store neurocomputational model that incorporates spreading activation processes.

Alternatively, the retrieval account by Whittlesea and Jacoby (1990) might be capable of predicting the findings above, which are better performance in related compared to unrelated trials and chunking of related but not unrelated stimulus-pairs that results in a zigzag pattern. Other compound-cue theories that do not assume that relatedness moderates the formation of compounds cannot easily explain the findings above. The theory by Ratcliff & McKoon (1988), which focuses on an enhanced familiarity for compounds containing related stimuli, might have problems accounting for the performance differences in the free recall used in the experiments described above. At best, the theory by M. S. Humphreys and colleagues (1993) could explain the better performance in the related compared to the unrelated condition in the free recall task. However, it remains unclear how the theory by Humphreys could explain the observed zigzag pattern. Because the retrieval account and an assumption of spreading activation processes (as in the dual-store neurocomputational model) predict actually similar effects and a similar “chunking” of related words for the task used by Davelaar and colleagues (2006), these theories might be to a certain degree compatible. Potentially, the dual-store neurocomputational model can explain precisely how the unitization that was assumed in the retrieval account could work.

Beneficial effects on the maintenance of evaluatively congruent stimuli

An attempt to rule out the influence of a consciously generated common category could be made by separating the features defining similarity and the features that have to be memorized. Measuring the effect of semantic overlap in working memory in a paradigm incorporating this feature into the design could provide profound insight into the existence of effects of automatic mutual facilitation of related concepts in working memory.

Unintentionally, this was achieved in an experiment by Jackson, Linden & Raymond (2014). Although they interpret their findings differently, their results are a prime example of a retroactive priming effect in working memory. In Experiment 3 of their study, their participants saw with either only angry or only happy faces at encoding and they were presented with a single neutral face at test. Most importantly, during the maintenance phase, a positive or negative word was presented. At test, participants indicated whether an old or a new identity was presented. Interestingly, they performed better in this task when valence of the encoded faces and word valence matched.[6] Therefore, their experiment provides evidence for a beneficial effect of evaluative congruency in working memory that originates during maintenance. Thus, the results

---

[6] Note that the authors interpret the 2 (face valence) × 2 (word valence) interaction as showing a congruence effect for angry faces but not for happy faces because the simple main effect of word valence for happy faces was non-significant. However, to decompose a 2 × 2 interaction in such designs is problematic because overall main effects (here for word valence) might cause an interaction pattern that seems to be asymmetric (i.e., here: a congruence effect for angry faces but not for a happy faces; see Wentura & Degner, 2010 for a discussion of this issue that is also relevant in evaluative priming research).

can again be explained by spreading activation theories and the dual-store neurocomputational model by Davelaar and colleagues (2006).

At first glance, an application of compound-cue theories to the result by Jackson and colleagues (2014) seems problematic. The compound-cue theory by M. S. Humphreys and colleagues (1993) can certainly not be applied to this finding: The assumption that the target face would be in the intersection of generated sets is rather implausible because the words should not lead to the generation of face-candidates in this change detection task. An explanation based on an enhanced familiarity by assuming that the face and the word would form a compound with a higher familiarity signal in congruent trials seems also implausible. The same critique holds for an explanation by the retrieval account by Whittlesea and Jacoby (1990). However, it should be noted that the formation of a compound (especially in congruent but not incongruent trials) cannot be completely ruled out. This is because a subset of the valent words used was related to faces (e.g. Simile, Frown, Laugh, Shout; or at least person or interaction related like Agree, Argue, Kiss, Kick). The possibility that faces and words formed a compound can therefore not be excluded completely. It should be also noted that, as the example stimuli make clear, there was not only a mere evaluative overlap between words and faces but also additional strong semantic and associative overlap between the two classes of stimuli. Be that as it may, their study provides compelling evidence that the maintenance of a stimulus can be boosted by related stimuli that are presented during maintenance, even if these stimuli stem from distinct classes and when they are only linked by semantic meaning. This backward directed effect on the maintenance of a related concept is one part of the mutual facilitation we assume as a working memory phenomenon.

In summary, there is promising evidence for both parallel activation and mutual facilitation in priming paradigms, as well as parallel activation and mutual facilitation in working memory. Most of these results can either be explained by spreading activation theories and a spread of activation back and forth between related concepts or by a subset of compound-cue theories. Potentially, both families of theories do not exclude each other. Better working memory performance in the case of relatedness as well as higher chunking can probably be explained within a single framework. Therefore, the theory by Davelaar and colleagues (2006), in which both effects were modeled and explained as originating from a spread of activation in a limited capacity short-term buffer provides a promising framework. Together with the three-process model for priming from Schmitz and Wentura (2012), nearly all discussed findings can be explained parsimoniously assuming mutual facilitation for concurrently maintained semantically related concepts.

1.5 Overview of Experiments

The overall goal of the empirical work conducted for this thesis was to add on to the notion of mutual facilitation of related concepts in priming and working memory. Therefore, different methods and tasks compared to previous studies were used, partially to avoid potential confounds of previous studies and partially to test our hypothesis differently to provide evidence for the generalizability of the effects. For example, while Schmitz and Wentura (2012) used *evaluative* overlap in a priming paradigm, we used *semantic* overlap in a priming task (Experiment 1a and 1b). While Davelaar and colleagues (2006) used free recall and *semantically* related stimuli, we manipulated *evaluative* congruency and we used the change detection task, which is another classical working memory paradigm (Experiments 2-4). If the same mechanism can explain the results by Schmitz and Wentura (2012) and Davelaar and colleagues (2014), the studies described below can contribute to bridging the gap between these two informative and contemporary lines of priming and working memory research. We tried to rule out several alternative explanations like strategic guessing or the use of common category retrieval cues and aimed for an investigation of automatic processes of relatedness.

For this purpose, Experiment 1a and 1b targeted parallel activation and mutual facilitation of semantically related concepts using a priming paradigm. In the two experiments, prime-target pairs were presented briefly and masked. A perceptual identification task was employed and we expected effects in accuracy. Stimuli were presented briefly and simultaneously with an SOA of 0ms. Experiment 1a was designed to validate the paradigm of using the perceptual identification task with SOA zero. With this experiment we aimed to provide first hints for parallel activation and mutual facilitation of simultaneously active related concepts in a semantic priming task. Experiment 1b was designed to provide more compelling evidence. In Experiment 1b, we merged a post-cue task with the perceptual identification task to measure semantic priming effects. Finding a priming effect with overall performance above 50% would provide evidence for parallel activation of both concepts and mutual facilitation in the case of relatedness. Furthermore, due to the application of the post-cue task, it could be shown that priming effects can emerge when the asymmetry between prime and target at encoding is strongly reduced. Finding a significant priming effect in this paradigm can most parsimoniously be explained by a mutual facilitation of semantically related concepts due to spreading activation or by a higher unitization of related concepts as assumed by the retrieval account from Whittlesea and Jacoby (1990). Furthermore, although a priming paradigm is used in Experiment 1b, a working memory component is incorporated: Using the post-cue task, participants have to store the encoded words and to maintain them until a post-cue indicates which word has to be named. Accordingly, it prepares the ground for the following experiments that used a classical working memory task. Besides the rationale mentioned above, evidence for parallel activation and mutual facilitation in the case of relatedness would also be of crucial importance because Schmitz and Wentura (2012) propose that their theory can be seen as contradictory to explanations of priming that rely on the classical sequence of target activation following prime activation. The use of the perceptual identification task further allows checking for influences of strategic guessing by investigating false responses.

In Experiment 2 (a-d), a working memory task was used: the change detection task. Faces showing emotional expressions were used as stimuli. Relatedness was manipulated by the overlap of the evaluative component. More precisely, faces in memory displays showed either all the same emotion (all showed happiness or all showed anger) in congruent trials or faces expressed different emotions (happiness and anger) in incongruent trials. In each of the four highly comparable experiments (2a-d), a set size of four was used. Besides the investigation of congruency effects with upright faces, we furthermore checked for effects induced by perceptual overlap: For inverted faces processing of the emotional component is reduced; however, the perceptual features stay the same as for upright faces (Fox et al., 2000). Therefore, finding no or a reduced congruency effect for inverted faces would indicate that a potential effect for upright faces is not merely based on perceptual overlap. Under this condition, a congruency effect for upright faces can be regarded as evidence for mutual facilitation of evaluatively congruent concepts. This can be explained by a spread of activation between congruent concepts. Finding a reversed effect would indicate interfering effects of evaluative congruency in working memory. However, based on the predictions derived from theories assuming a spread of activation, positive effects should be expected. Another purpose of Experiment 2a-d, which is less related to the overall goal of this thesis but nonetheless an important aspect of the study itself, was to replicate the angry face benefit (i.e., better performance for angry compared to happy or neutral faces, see Jackson et al., 2009).

In a further series of experiments, that is, Experiment 3a-c, a similar task to Experiment 2a-d was implemented, but set size two was chosen. Again evaluative congruency was manipulated. To achieve task-difficulty comparable to Experiment 2a-d despite the reduced memory load, stimuli were presented at different locations at test compared to encoding. Basically, the same effects could be assumed as for Experiment 2a-d, better performance in congruent compared to incongruent trials, which can be explained by a spread of activation. However, other influences on performance might have unintentionally been introduced by changing locations of the stimuli from study to test. These potential influences are discussed when the experiment is described in more detail.

In Experiment 4, we used a change detection task in which the stimuli were presented at the same positions at encoding and at test (as in Experiment 2a-d) to avoid any influence of a location change at test. Again, evaluative congruency was manipulated. As in Experiment 3a-c, set size two was used to avoid even a slight evaluative congruency in the incongruent condition, which would arise when four faces are used. We predicted finding an advantage in congruent compared to incongruent trials, as spreading activation theories suggest (J. R. Anderson, 1976, 1983, 1993; Collins & Loftus, 1975; Davelaar et al., 2006). The paradigm we used was roughly similar to the procedure used in the Experiment 3a-c, except that the spatial arrangement did not change from study to test. However, to gain further insight into the underlying mechanism of the congruency effect, we incorporated an additional manipulation into the design. One face was marked as task-relevant. Further, besides the target face, the task-irrelevant face could also be replaced by another face. Accordingly, we were able to investigate the effect of such an irrelevant change. This enabled us to test the assumption of mutual facilitation due to spreading activation processes against an alternative explanation based on the compound-cue theory by Whittlesea and Jacoby (1990), which can predict a higher effect of irrelevant changes in congruent compared to incongruent trials.

2 Parallel activation and mutual facilitation in semantic priming[7]

The goal of these first experiments was to provide evidence for the notion of mutual facilitation of simultaneously active semantically related concepts in priming. This can be predicted by most theories described above, first and foremost by the three-process model by Schmitz and Wentura (2012) and by spreading activation theories (J. R. Anderson, 1976, 1983, 1993; Collins & Loftus, 1975) including the dual-store memory model by Davelaar and colleagues (2006) that also incorporates a spreading activation process. Because Schmitz and Wentura (2012) provide evidence for these claims using evaluative priming, a logical next step is the test of their assumption in semantic priming, for which the same principles should apply. Compelling evidence for automatic mutual facilitation in priming should furthermore rule out or control for strategic processes.

Evidence for the assumption of parallel activation and mutual facilitation in priming is relevant from two perspectives. First of all, within the framework of this thesis and its focus on working memory research, providing evidence for parallel activation of concepts in semantic priming can create a link to working memory research (see Schmitz & Wentura, 2012). By applying a priming task that gains similarity to memory tasks (Experiment 1b), observing evidence for mutual facilitation implies that the investigation of mutual facilitation in working memory is a worthwhile endeavor. Therefore, one aim of the first two experiments in this thesis is to pave the way for subsequent studies investigating mutual facilitation in working memory.

---

[7] This chapter is mostly identical to Scherer and Wentura (2018). However, it was partly shortened to avoid large-scale redundancies. Furthermore, the data of Experiment 1a are also reported in my master thesis (Scherer, 2015). Thus, planning and conducting this experiment was pre-dissertation work. Notwithstanding, Experiment 1a is also reported here because of its close relatedness to Experiment 1b and their joint report in a paper by us (Scherer & Wentura, 2018). In contrast to the aforementioned master thesis, the following section takes new theoretical foci into account and additionally provides a more complex common discussion of Experiment 1a and Experiment 1b, which considers theories beyond those that were mentioned in my master thesis.

The second reason why parallel activation and mutual facilitation in semantic priming are very intriguing is that they are arguably controversial assumptions, although both processes are highly plausible. They can be derived from several theories and they are supported by several current and classical findings as outlined earlier. However, parallel activation and mutual facilitation are arguably controversial assumptions because in most classical semantic priming studies, the target has to be categorized as a word (versus non-word) or it has to be named as soon as possible and the prime is presented briefly and it *precedes* the target (for reviews, see McNamara, 2005, 2013; Neely, 1991). At first glance, this might make it unnecessary to assume that the stimuli presented one after the other can be active at the same time. Interestingly, the notion of parallel activation can be regarded as puzzling because it can arguably be more easily explained by some theories of semantic priming than by others. As already discussed, besides the definition of semantic priming also the label "prime" for the preceding stimulus alludes to the usual temporal order of the stimuli. This prototypical feature of semantic priming experiments invites theoretical explanations of semantic priming that presuppose an asymmetry of prime and target: These explanations focus on the encoding of the target, which is facilitated by the processing of a related prime. When the sequence of the prime activation preceding the activation of the target is mandatory for a model to predict priming (Masson, 1991, 1995), explaining evidence for parallel activation and mutual facilitation in the case of relatedness becomes a challenging endeavor for these models (see Schmitz & Wentura, 2012; Schmitz et al., 2014). Thus, these theories need some additional assumptions to account for evidence for parallel activation and mutual facilitation. Following these considerations, evidence for parallel activation and mutual facilitation can contribute to establishing these two assumptions as criteria which semantic priming theories should be able to account for (we will postpone further discussion of this issue to the Discussion of the first two experiments to review the different theories directly in light of our results).

2.1 Combining the post-cue task and the perceptual identification task (Experiment 1a and 1b)

As described above, for concepts being activated simultaneously some recent theories and evidence suggest mutual facilitation when the concepts overlap semantically. Therefore, we aimed for providing evidence for parallel activation and mutual facilitation of related prime-target pairs as described by Schmitz and Wentura (2012) using a semantic priming paradigm. They suggest that parallel activation describes a simultaneously active prime and target. The process of a mutual facilitation constitutes of two parts. On the one hand, a prime facilitates the activation of a related target concept, and on the other hand, the target helps to maintain the prime when both stimuli are related. To test for mutual facilitation of related concepts, a perceptual identification task was employed, presenting prime-target pairs briefly and masked, with an SOA of 0 ms (i.e., prime and target were presented concurrently, one above the other). Participants were instructed to identify the target. In Experiment 1a, a cue defining the target was presented at stimulus onset, whereas in Experiment 1b the cue was not presented before the offset of stimuli. Accordingly, in Experiment 1b, a post-cue task was merged with the perceptual identification task. This feature enriches the priming task validated in Experiment 1a with a memory component. It further takes away the asymmetry between prime and target during encoding. We expected significant semantic priming effects in both experiments, a result that is compatible with the view that two concepts can both be activated in parallel and that they can mutually facilitate each other if they are related.

2.1.1 Paradigm and overview

The literature on semantic priming with its experimental varieties provides several suggestions for a paradigm suited to test our hypotheses. First, a parallel presentation of prime and target suggests itself to avoid an asymmetry between prime and target at encoding. There are indeed several studies with an SOA of 0 ms (see, e.g., J. E. Anderson & Holcomb, 1995; de Groot, 1984; Masson, 1991; Meyer & Schvaneveldt, 1971) that found semantic priming effects. This alone, however, is not sufficient to provide evidence for our hypothesis because participants could still process the prime first and the target second.

Therefore, second, the asymmetry of prime and target should be removed, that is, both items should be candidates for identification and a cue determining the to-be-reported stimulus should not be presented before the onset of the items. Indeed, a post-cue priming task was already used a few times in semantic priming research (see, e.g., Dallas & Merikle, 1976a, 1976b; G. W. Humphreys, Lloyd-Jones, & Fias, 1995; Murphy, 2010; Murphy & Green, 2011). Usually, semantic relatedness of the two words leads to faster naming latencies (but see G. W. Humphreys et al., 1995). However, most of the previous post-cue experiments had rather long presentation times and/or late onset of the cue. This might have induced expectancy-based effects similar to those found with long SOAs in standard priming tasks (see C. A. Becker, 1980).[8]

Third, a perceptual identification task (i.e., a briefly presented stimulus overwritten by a mask has to be identified; Evett & Humphreys, 1981; Pecher, Zeelenberg, & Raaijmakers, 2002) seems to be a good choice. Responses are based on the identification of a given stimulus which makes it plausible that the representation of this stimulus was active while the participant responded. Moreover, binary decision tasks (e.g., lexical decision) which are often used in semantic priming research are prone to post-lexical strategies (Balota & Lorch, 1986; de Groot, Thomassen, & Hudson, 1982; de Groot et al., 1986; Neely & Keefe, 1989; Seidenberg et al., 1984) which might make results equivocal. Finally, the analysis of erroneous responses in the perceptual identification task allows for a test of strategic guessing.

Of course, the perceptual identification task resembles the often-used naming task (i.e., the target word, which is clearly perceptible, has to be named as soon as

---

[8] Dallas & Merikle (1976b) used short presentation durations *and* a short post-cue delay. However, some features did not conform to standard practice: The stimulus lists were not counterbalanced and each target was presented six times throughout the experiment, whereas each prime was presented only twice.

possible). However, there are clear differences. Methodologically, the major difference between the standard naming task and the perceptual identification task is that the stimuli are masked in the latter task. Therefore, identification is more difficult, leading to a focus on accuracy instead of reaction times. With regard to the underlying processes, the perceptual identification task relies more on semantic processing than the naming task. The short presentation duration prevents responses based on a non-lexical route (i.e., on a grapheme-phoneme rule system; see, e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) that can by-pass the semantic system. Furthermore and most importantly, perceptual identification allows to check more directly for expectancy based processes (because these processes potentially lead to guessing of a prime-related but not presented concept). Therefore, in our main experiment (Experiment 1b) we used an amalgamation of the perceptual identification task and the post-cue paradigm, thereby employing an SOA of 0ms. We start, however, by Experiment 1a, that is, an adaptation of the perceptual identification task to a presentation with SOA=0ms. This modification not only allows for the replication of the semantic priming effect in a perceptual identification task with SOA = 0 ms but also paves the way for the Experiment 1b, where we cued the to-be-identified stimulus after the offset of the stimulus pair.

2.2 Experiment 1a

2.2.1 Method

Participants. Twenty-one undergraduate students (18 females; age range 19-42 years, Md=20 years) participated for partial course credit. All of them were native speakers of German and had normal or corrected-to-normal vision.[9]

A power analysis using GPower (Faul, Erdfelder, Lang, & Buchner, 2007), assuming an effect size of d_Z = .77—the effect obtained in Experiment 1B by Pecher and

[9] There were two further participants who completed the task to obtain their course credit but did not fulfill our a priori criteria for inclusion, one non-native speaker of German and one participant with dyslexia.

colleagues (2002)—, α = .05 (two-tailed), and a sample size of N = 21, returned an estimated power of 1-β = .96.

Design. A one-factorial within-participants design was used with the sole factor of semantic relatedness (related vs. unrelated). A between-participants control factor was added to control for the effects of counterbalancing the word lists (see *Materials* section).

Materials. We selected 120 associated word pairs for the main phase of the experiment from Hager and Hasselhorn (1994) and Russell (1970). Although stimuli from association norms were used, the word-pairs were – as a rule – semantically related as well. The list of the associated prime-target pairs was randomly split into two subsets (A and B), each consisting of 60 pairs. Then two subsets of unrelated prime-target pairs (A' and B') were created by re-pairing the associated word pairs within each subset A and B. Specifically, in the unrelated condition, each prime was paired with an unrelated target of the same word length as the corresponding related target. Half the participants were presented with the related pairs of Set A and the unrelated pairs of Set B; for the remaining participants, the assignment was reversed. Another 60 unrelated word pairs were selected for the calibration phase of the experiment. Finally, 12 additional words were used for the practice trials.

Procedure. Participants were seated in front of a 15-in. CRT screen at a distance of approximately 0.6 m. The trials within the practice phase, the calibration phase, and the experimental trials were virtually the same (see Figure *1*): Participants first saw a fixation cross, which was followed by a forward mask, consisting of two sequences of 13 @-symbols each. The masks were replaced by the two word stimuli (i.e., prime, target). Words were presented one above the other and centrally aligned. To the left and right of each word, @-signs were added to obtain a string length of 13 characters. Both words together covered a visual angle of approximately 11.42 × 2.96°. A single word covered approximately 11.42 × 1.15°.

The target was marked with arrows on the left and right side of the string, appearing simultaneously with the onset of the words. In half the trials, the target was presented in the upper position; in the other half, it was displayed at the lower position. Following the presentation of prime and target (for the duration, see below), a backward mask (again two sequences of 13 @-signs) replaced the word stimuli. Subsequently, three question marks appeared in the center of the screen and disappeared after a response was given and coded by the experimenter, who saw the correct response on a second screen.

Participants were instructed to name the target word as soon as possible. If participants did not recognize the target, they were requested to respond "*null"*. After the response was coded by the experimenter, a feedback message was presented for 1000 ms ("*correct", "false",* or "…" [in the case of a *null* response]). Participants naming the prime instead of the target were reminded to name the target by the message "*Please name the marked word"* (for 2000 ms).

Participants first completed six practice trials with a prime-target display of 800 ms duration. During the calibration phase, the presentation times for the prime-target displays were 110, 120, 130, 140, 150, and 160 ms (in random order). Each of these presentation times was used 10 times. The performance was used to calculate an individual presentation time for the main phase. The stimulus duration allowing to correctly identify the target in 60% of the trials was estimated with logistic regression (see Pecher et al., 2002). Above-chance performance (> 50%) is necessary for the following reason: If performance is below 50%, one can argue that participants always focused on one location. If by chance this is the prime (~50% of the trials), the response will never be correct. If by chance the focus is on the target (again ~50% of the trials), the trial mimics a standard priming task: the target will be the dominantly processed and sustained stimulus; transient processing of the prime might cause encoding facilitation of the target if it is related. Most participants (n = 15) required a presentation duration of 160 ms to achieve the 60-%-criterion (overall mean duration was M = 151ms). For trials of the main phase, the individually determined presentation time was used. The main phase comprised 120 trials.

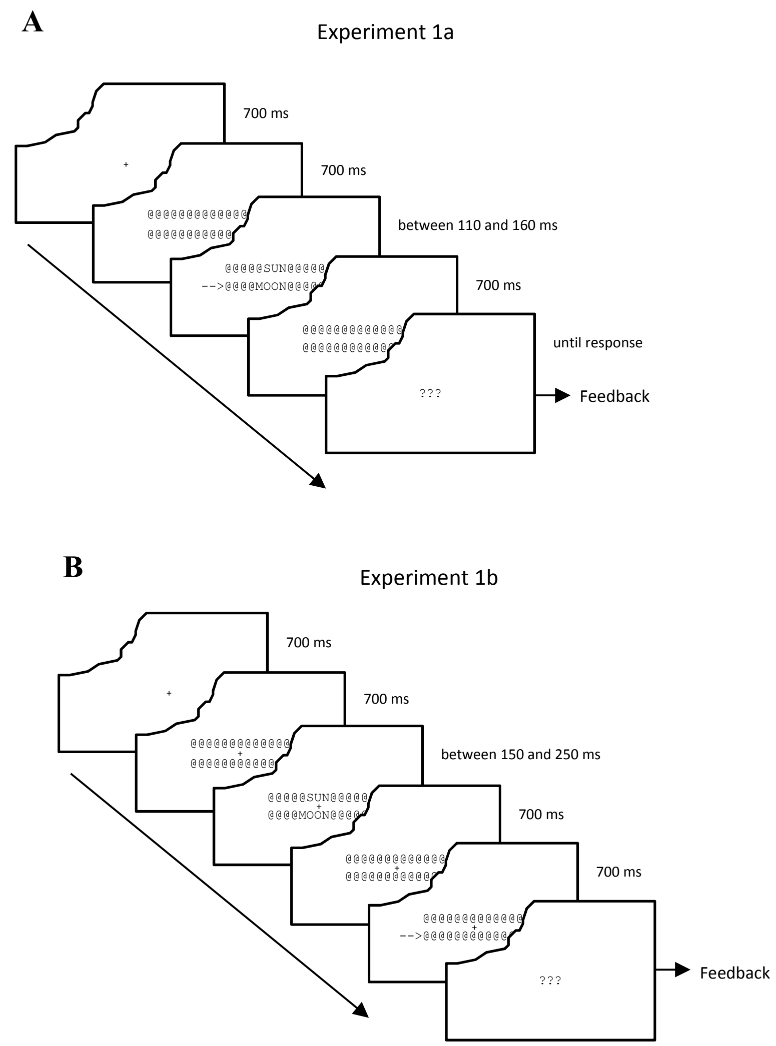


Figure 1. *Trial sequence of the perceptual priming task in Experiment 1a and Experiment 1b.*

At the end of the experiment, participants filled in a short questionnaire. This measured task difficulty and comprehensibility of instructions and participants were additionally prompted to report the strategies they applied during the experiment (in a free response format).

2.2.2 Results

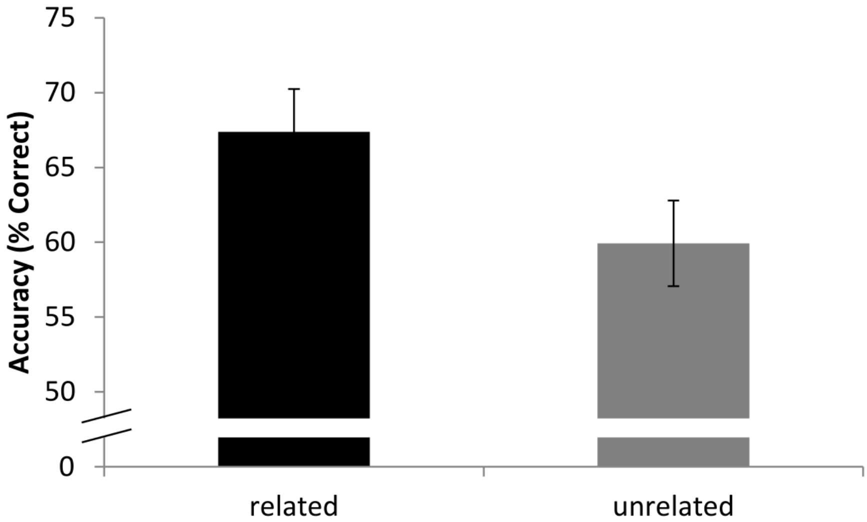


Figure 2 shows the mean percentages of correct target identification. Participants identified significantly more target words if these were accompanied by related primes compared to unrelated primes, M_{Δ} = 7.46% (*SD* = 8.88%; 95% CI [3.41-11.51]), $F(1,19) = 19.26$, $p < .001$, $\eta_p^2 = .503$.[10] In both the related and the unrelated condition, performance was significantly above 50%, with *M* = 67.38% (*SD* = 17.22) in the related condition, $F(1,19) = 20.29$, $p < .001$, $\eta_p^2 = .516$ and *M* = 59.92% (*SD* = 15.45) in the unrelated condition, $F(1,19) = 8.88$, $p = .008$, $\eta_p^2 = .318$.

Further analyses revealed that only four participants had responded, in only one to three unrelated trials per participant, with the word that would have been the target in the related condition. Thus, evidence for strategic guessing was rather scarce. Discarding these participants did not affect the outcome; there was still a significant priming effect of M_{Δ} = 8.53% (*SD* = 9.52%; 95% CI [3.64-13.42]), $F(1,15) = 14.30$, $p = .002$,

[10] In all analyses, we added counterbalancing group as a between-participants control factor in order to use the correct error term (see Pollatsek & Well, 1995) and to ensure that results were not affected by (slight) differences in group size. Results were essentially the same without adding this factor. For the sake of brevity, we will not report the results for this factor.

$\eta_p^2 = .488$.

In the post-experimental questionnaire, four participants reported using the strategy of guessing associated words upon identifying the non-marked word. Discarding these participants likewise did not affect the outcome; there was still a significant priming effect of $M_\Delta = 5.39\%$ ($SD = 7.67\%$; 95% CI [1.45-9.33], $F(1,15) = 10.59$, $p = .005$, $\eta_p^2 = .414$.

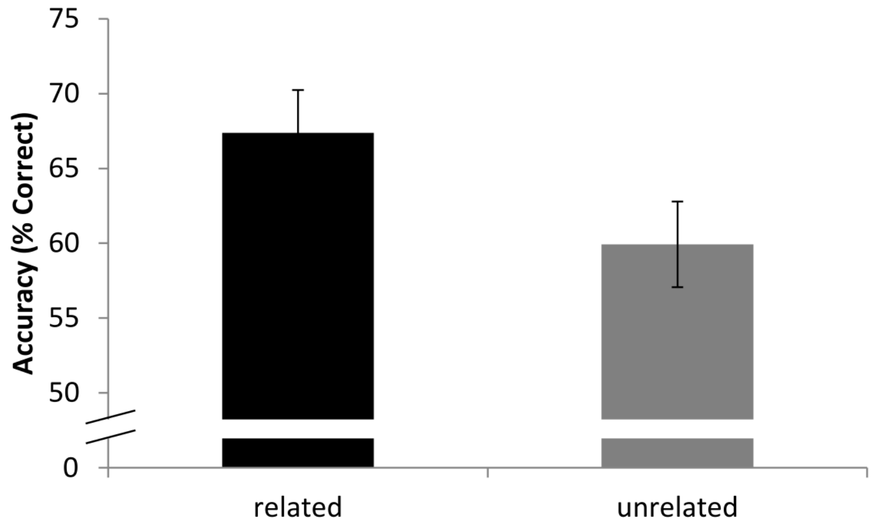


Figure 2. *Means for the related and unrelated trials (Experiment 1a). Error bars are 95% within participants confidence intervals (Loftus & Masson, 1994).*

2.2.3 Interim discussion

In Experiment 1a, we found a marked semantic priming effect using a perceptual identification paradigm, comparable to the results by Evett and Humphreys (1981) as well as Pecher and colleagues (2002), however with an SOA of 0 ms. The effect held if participants who showed some evidence of guessing, or participants who claimed to have done so in the post-experimental questionnaire, were excluded. Moreover, the priming effect was accompanied by an overall performance level greater than 50%. This constitutes preliminary evidence for the parallel activation of prime and target. However, a priming effect for trials with a backward cue would constitute more decisive evidence:

If participants learn which word has to be named only after the offset of prime and target stimuli, the argument of parallel activation could be made more forcefully.

As usual with the perceptual identification task (see, e.g. Evett & Humphreys, 1981; Pecher et al., 2002), we did not record response times because the rather lower number of correct trials makes the validity of response times rather questionable from the start on (since response time analyses are restricted to correct trials). Nevertheless, one might see a caveat of Experiment 1a in the sole measurement of accuracy because it cannot be ruled out that the effect can alternatively be explained by a speed-accuracy tradeoff. Therefore, in Experiment 1b, response times were additionally recorded.

2.3 Experiment 1b

2.3.1 Method

Participants. A total of 30 undergraduate students (27 females; age range 18-31 years, Md=21 years) participated for partial course credit. All of them were native speakers of German and had normal or corrected-to-normal vision.

A power analysis based on the effect size of Experiment 1a ($d_Z = .84$), $\alpha = .05$ (two-tailed), and $N = 30$ yielded a power of $1\text{-}\beta = 99.8\%$. Alternatively, with $N = 30$ and $\alpha = .05$ (two-tailed), an effect of $d_Z = .53$ (i.e., a medium-sized effect; Cohen, 1988) could be detected with $1\text{-}\beta = .80$.

Design, Materials, and Procedure. Everything was the same as in Experiment 1a, with the following exceptions. The main difference was that the target was now marked at offset of the words. In detail, the backward mask (see Figure 1) was prolonged from 700ms to 1400ms; after 700ms, the target-defining arrow cue was presented at either the top or the bottom location. We now additionally recorded response latencies, registering responses by a voice key.

A minor difference to Experiment 1a concerns the calibration phase. Again, six different presentation times were used. However, presentation times were between 150 and 250 ms (varied in 20 ms steps), given that the majority of participants in Experiment 1a required the longest presentation duration (i.e., 160 ms). For the main part of the experiment, there were eleven possible presentation times, namely all 10ms steps between 150 and 250 ms. Again, an individual presentation time allowing to correctly identify the target in 60% of trials was estimated with logistic regression. Mean presentation duration was $M = 208$ ms ($SD = 47$ ms).[11]

We used a higher screen resolution compared to Experiment 1a, but the same font for the words. Consequently, the words covered a smaller visual angle on the screen compared to Experiment 1a. Both stimuli together with their surrounding @-signs covered approximately 9.15 × 2.48° visual angle (9.15 × 0.96° visual angle for a single word). This was done to reduce the probability that participants preferentially attended to a single word.

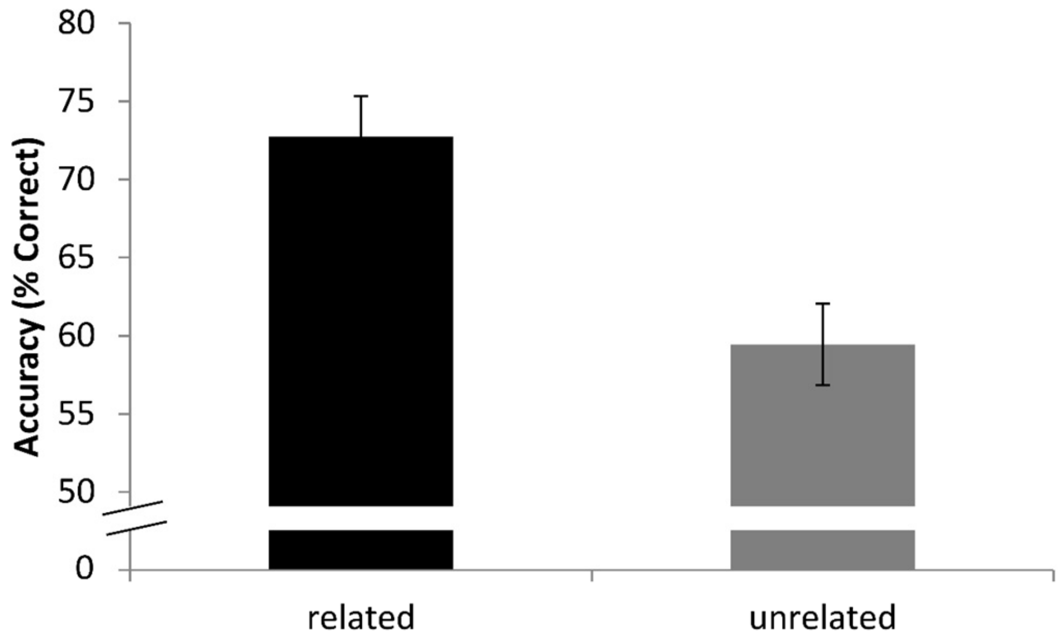
For the same reason, a small fixation cross in the middle of the screen between the word positions was added to the display. Participants were instructed to fixate it and not to shift attention to a single word. To encourage compliance, we added supplementary trials with a different task. In 20 trials of the main part (2 trials during practice; 10 trials during calibration), the fixation cross was replaced by the letter *X*. Participants were instructed to name the letter as soon as possible.

2.3.2 Results

Accuracy. Figure 3 shows the mean percentages of correct target identification. Participants identified significantly more targets in the related condition compared to the unrelated condition, $M_{\Delta} = 13.28\%$, ($SD = 9.86$), 95% CI = [9.6-16.96], $F(1,28) = 54.40$, $p < .001$, $\eta_p^2 = .652$. The performance in both the related and the unrelated condition

[11] It should be noted that there were two modal values, 150ms ($n = 11$) and 250ms ($n = 14$). Results were not moderated by a between-participants factor presentation time (median split).

were significantly above 50%, indicating that more than a single concept is maintained per trial, M = 72.72 (SD = 14.30), $F(1,28) = 73.85$, $p < .001$, $\eta_p^2 = .725$, for the related condition; M = 59.44 (SD =15.92), $F(1,28) = 12.01$, $p = .002$, $\eta_p^2 = .300$, for the unrelated condition. Comparing the effect of relatedness of both experiments revealed that the benefit for related compared to unrelated trials was more pronounced in Experiment 1b (M_Δ = 13.28%, SD = 9.86) than in Experiment 1a (M_Δ = 7.46%, SD = 8.88); $F(1,47) = 6.57$, $p = .014$, $\eta_p^2 = .123$.



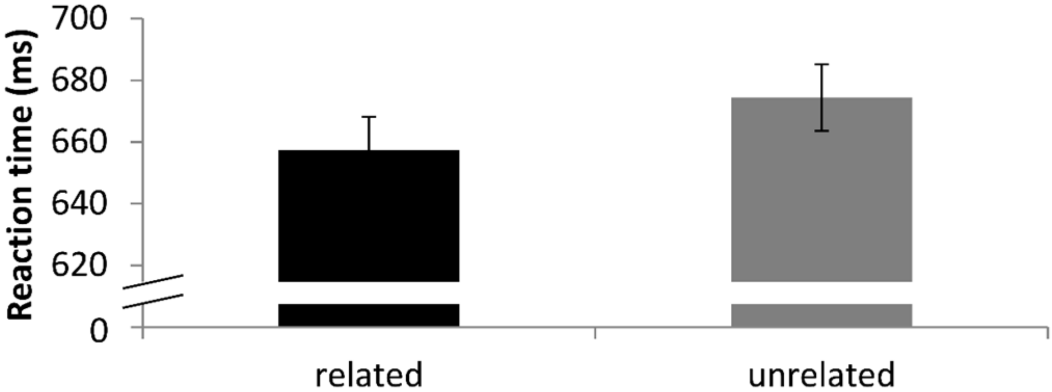


Figure 3. *Mean percentage correct (top) and mean response times (in ms, bottom; measured from onset of cue) in Experiment 1b for related and unrelated trials (Error bars show 95% within-participants confidence intervals, as described by Loftus & Masson, 1994).*

Analyses of erroneous responses in unrelated trials revealed that $n = 16$ participants, on a total of 17 trials (0.9%), responded with the target word of the related condition. Thus, evidence for prime-triggered guessing was modest. Nevertheless, adding strategy use as a between-participants factor, the priming effect remained significant, $F(1,26) = 63.93$, $p < .001$, $\eta_p^2 = .711$ and was not moderated by strategy use, $F < 1$ ($F < 1$ for the main effect of strategy use). Constraining the analysis to those participants who showed no evidence of strategy use still yielded a significant priming effect, $M = 13.81\%$, ($SD = 9.14$), 95% CI [8.53-19.09], $F(1,12) = 40.51$, $p < .001$, $\eta_p^2 = .771$.

As a second check, we asked participants explicitly whether they had used the strategy of guessing the target on the basis of the other word if they were not able to detect it. A total of $n = 17$ participants acknowledged the use of such a strategy. Adding strategy use as a between-participants factor did not result in an interaction relatedness × strategy use, $F(1,25) = 1.82$, $p = .189$, $\eta_p^2 = .068$. The priming effect still remained significant, $F(1,25) = 78.84$, $p < .001$, $\eta_p^2 = .759$ ($F < 1$ for the main effect of strategy use). A significant priming effect was observed even when analysis was restricted to the 13 non-strategy users, $M = 72.18$ ($SD = 16.46$) for related trials, $M = 56.41$ ($SD = 18.66$) for unrelated trials, $F(1,11) = 38.01$, $p < .001$, $\eta_p^2 = .776$.

Response times. Analyses of response times (RT) were restricted to trials with correct responses. RTs that were 1.5 interquartile ranges above the third quartile of the individual RT distribution (Tukey, 1977) or were below 200ms, as well as RTs of trials in which participants made a sound prior to their response (e.g., clearing their throat), were discarded (8.05% of trials). Figure 3 (bottom) shows mean RTs for the conditions of interest. We observed significantly faster response latencies in related trials compared to unrelated trials, $F(1,28) = 5.04$, $p = .033$, $\eta_p^2 = .152$. Accordingly, the reaction times showed the same pattern as the accuracy data, providing evidence that there was no speed-accuracy trade-off.

2.3.3 Interim discussion

The results of Experiment 1b clearly supported our hypothesis. Semantic priming effects were obtained without the target being defined at its encoding. The observed effect was not moderated by strategic influences and it was found in the subsamples that showed no evidence of strategy use. Finally, it was also demonstrated that the effect in accuracy was not caused by a speed-accuracy trade-off, since (correct) response times revealed the priming effect as well. Therefore, we again provide evidence for the applicability of the perceptual identification task in semantic priming (see also Evett & Humphreys, 1981; Pecher et al., 2002). Additionally, our result clearly indicates that using a post-cue paradigm, the effect does not vanish as it could be speculated due to the insignificant effect in a closely similar study by G. W. Humphreys and colleagues (G. W. Humphreys et al., 1995). Instead, the effect is even more pronounced than in Experiment 1a. There can be several reasons for this. First, a working memory component is incorporated into the design by marking the target only after the offset of prime and target. If a mutual facilitation can take place in working memory, as suggested by Schmitz and Wentura (2012) as well as Davelaar and colleagues (Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Davelaar et al., 2006 see also Haarmann & Usher, 2001; Usher & Cohen, 1999) it is plausible to assume that effects of relatedness are increased when a working memory component is introduced into a priming task. Furthermore, spreading activation theories can predict a monotonic increase in the size of the relatedness effect as the SOA between the onset of the two-word display and the onset of the target-cue increases. This is because the received activation at the target node should increment when time proceeds, resulting in a build-up of activation. However, further research is needed to validate this hypothesis.

2.4 Discussion

Our results show that semantic priming effects can be found with the combination of the perceptual identification task with a post-cue, using an SOA of 0ms. Thus, the most straightforward explanation is that both words had to be encoded and distinct representations had to be maintained until the location of the to-be-identified stimulus was marked. The priming effect can be understood as a mutual facilitation of prime and target activation, as assumed by Schmitz and Wentura (2012). Before discussing how our results can be reconciled with prevailing theories of semantic priming, we will list limitations of the current study and compare the current experiments to similar other tasks.

2.4.1 Limitations

First, we cannot rule out a speed-accuracy trade-off for Experiment 1a, in which only accuracy was measured (in accordance with other studies using the perceptual identification task; Evett & Humphreys, 1981; Pecher et al., 2002). Nevertheless, for Experiment 1b, in which a similar procedure was used, reaction times showed the same pattern as the accuracy data. Please note that accuracy was rather low and analysis of reaction times is only based on response times of correct trials. Therefore, effects in response times are in the current study less meaningful then effects in accuracy.

Another limitation concerns whether the SOA of 0ms does truly avoid initial asymmetric processing of the prime and the target or not. Please note, that for Experiment 1b the focus is on a mutual facilitation of concurrently active concepts, which could take place in working memory. Therefore, the initial sequence of processing for both words is less relevant and evidence for a parallel activation in Experiment 1b is not affected by critique in this way. To additionally ensure that prime and target are also *initially* processed simultaneously we implemented several features. First, in both experiments, the two words were presented simultaneously to enable mutual facilitation to accumulate from the onset on, contrasting with the dominant approach of sequential prime target presentation. Second, until the onset of both words, a participant cannot know which of the two words will be the target. Both points should slow down the initiation of an anticipatory strategy that aims to guess the meaning of the target from the prime and therefore make such a strategy less effective. Third, in Experiment 1b, the target was not marked before the offset of both words, further removing the asymmetry between prime and target. Nonetheless, these measures might not fully remove any asymmetry between prime and target. As targets we chose words that were generated as associates to the words that we used as primes, using material from word norms. Therefore, association strength from the prime to the target may not be the same as the association strength form the target to the prime. This could bias a mutual facilitation more into one direction than the other. Whether this difference is processed by the participants in that way that it leads to a preference in the initial orientation of attention to one stimulus or the other remains unclear.

2.4.2 The relationship of the present results to previous research

The observed effects are in accordance with classical research on semantic priming, indicating a positively signed effect of semantic relatedness. The current tasks replicate the semantic priming effects obtained in the perceptual identification task by Pecher and colleagues (2002) as well as Evett and Humphreys (1981).

Despite being in line with previous research on priming, the current data of Experiment 1b reveal a rather different finding compared with some studies using post-cue tasks. First, studies in which related and unrelated picture-stimuli instead of words were used indicate interference due to semantic relatedness. However, for picture naming, there is evidence that the object-attribute integration (for example color-form integration) is the source of these interference effects (see Dean, Bub, & Masson, 2001; Hocking, McMahon, & de Zubicaray, 2010). This rationale cannot be applied to our experiments. In Experiment 3 by G. W. Humphreys and colleagues (1995), colored words and color cues were used and neither positively signed priming nor interference effects were observed. There, still, object-attribute integration could have influenced the results with the word being the "object" and color defining the attribute. In another post-cue task by Dallas and Merikle (1976a, 1976b) with words as stimuli and post-cues more similar to ours, faster naming latencies for semantically related words were observed. As in our design, it can be assumed that object-attribute integration did not influence their results and therefore a positively signed effect of semantic relatedness was observed. In this regard, it could be relevant whether the target is defined by rather peripheral features or not. Alternatively, differences in findings obtained by different versions of the post-cue task could be explained by variations in task difficulty. For Experiment 1b of the current study, not only using perceptual identification instead of naming introduces a high task difficulty, but also the use of peripheral arrows to mark the target might have this effect. Therefore, further research is needed to investigate which factors interact with the effect of semantic relatedness in post-cue tasks. But given the beneficial effect of semantic relatedness in our task, how can we reconcile our results with prevailing theories of semantic priming?

2.4.3 Semantic network models

As was clear from the outset, spreading activation theories (e.g., semantic network models) can easily account for the results. According to these theories, both words activate their corresponding node, and, for related words, spreading activation between these nodes contributes to a mutual maintenance of this activation. For unrelated concepts, the spread of activation for each concept would not reach the distant part of the network in which the other concept is active. Therefore, related but not unrelated concepts can mutually maintain their activation and spreading activation theories can easily account for the current data. However, these models have been criticized within the domain of priming research due to some incommensurate results (e.g., Bodner & Masson, 2003; Masson, 1995; Whittlesea & Jacoby, 1990).[12] They also have been

[12] For example, effects of the relatedness proportion on non-strategic priming (Bodner & Masson, 2003) cannot be explained by classic spreading activation theories. Another critical

criticized more generally, among other factors e.g. (Johnson-Laird, Herrmann, & Chaffin, 1984), for their amodal symbolic representation of concepts that does not account for evidence that perceptual simulations can underlie the processing of meaning (e.g. Johnson-Laird et al., 1984; Pecher, Zeelenberg, & Barsalou, 2003). In contrast to that, a theory of semantic priming should ideally be a plausible model of memory in general.

2.4.4 Parallel-distributed models

As a general framework of memory, parallel-distributed theories are more attractive because of their capability to learn and because they seem better suited to model modality-specific features. Basically, priming effects are attributed to the similarity of related primes and targets in semantic space, compared to unrelated stimulus pairs (Cree, McRae, & McNorgan, 1999; Masson, 1991, 1995; McRae, de Sa, & Seidenberg, 1997; Moss, Hare, Day, & Tyler, 1994; Plaut, 1995; Plaut & Booth, 2000; Sharkey & Sharkey, 1992). Related concepts have similar activation patterns; unrelated concepts barely overlap in their activation. If the target is presented after the prime, the initial activation pattern representing the prime is updated to represent the target. Accordingly, due to the overlap, the activation pattern of the prime can transform quickly into the activation pattern of the target when prime and target are related. If there is no overlap, as in the case of an unrelated prime-target pair, the conversion is slower. Accordingly, these models necessarily assume a temporal order of the activation of the two concepts and distinct identification of the target at the time of its encoding. Thus, without additional

---

point for spreading activation theories is that intervening unrelated words presented between prime and target can reduce priming effects (Masson, 1995). According to spreading activation theories with localist representations, the spreading activation process in one part of the network should not be influenced by an intervening unrelated stimulus that is activated in a distant part of the network. Furthermore, a finding by Whittlesea and Jacoby (1990) is challenging for spreading activation theories. The naming latency to a repetition of the prime (which is presented after the target) is faster when the target is degraded and related to the prime. However, a backward directed influence which is fundamentally altered by degradation of the target word should not be assumed when an automatic spread of activation can explain priming effects.

assumptions, these models cannot account for parallel activation of concepts leading to mutual facilitation in the case of relatedness. However, this potential difficulty can be overcome in two ways.

First, the functions of (a) parallel maintenance of two (or more) items and (b) mutual facilitation might be provided by two different sub-systems. For example, a working memory module might hold the two lexical entries corresponding to prime and target; a semantic distributed memory system is accessed by the lexical entries but is constrained to one active pattern at a given point in time. If the active pattern accidentally corresponds to the target when the cue appears, the target will be easily named, irrespective of whether the prime is related or unrelated. If, however, the active pattern accidentally corresponds to the prime when the cue appears, the pattern has to change to the target pattern (if we assume for a moment that semantic access is needed for the task). In these cases, the relatedness can be effective as proposed by the parallel-distributed models of priming. Thus, our results can be accommodated by the prevailing models, however, at the price of structural constraints.

Second, one might think of a refinement of these models such that they can account for (a) parallel activation of several concepts and (b) mutual facilitation in the case of semantic relatedness. To solve point (a), some (working) memory models already incorporate synchronous firing of feature nodes that belong to the same concept (e.g., Raffone & Van Leeuwen, 2001; Raffone & Wolters, 2001; Vogel, Woodman, & Luck, 2001; Wolters & Raffone, 2008). Feature overlap of simultaneously maintained concepts is implemented by nodes that alternate their synchronization between patterns (Raffone & Van Leeuwen, 2003). Point (b), that is, the case of semantic priming, is yet to be addressed.

2.4.5 Memory-based accounts

Compound-cue theories (Dosher & Rosedale, 1989; Ratcliff & McKoon, 1988) can model a parallel activation of prime and target concepts in semantic priming paradigms. Generally, when compound-cue models are applied to priming, it is assumed that there is a memory-cue containing the target item as well as elements of its context, which can include the prime. Therefore, prime and target can be active in parallel. However, these theories dominantly explain priming effects found in binary decision tasks (e.g., lexical decision). Thereby, compound-cue theories need modification or further specification to be applied to naming tasks and perceptual identification (McNamara, 2005). An exception is the model by M. S. Humphreys, Wiles and Bain (1993), who provide a way to apply the compound-cue theory to the naming task in semantic priming. Their model predicts that when there are two cues to generate a response, each cue can generate an associative set. If a specific candidate is in the *intersection* of the generated associative sets of both cues (and the number of candidates in the intersection is rather small), it is likely that this candidate will be reported. For example, the cue *a mythical being* and the cue *rhymes with post* can be combined to generate a specific candidate in the intersection of both sets generated by the single cues. Therefore, the word *ghost* is generated with a high probability. The semantic priming effect in the naming task can be explained similarly. When there is an association between the prime and the target, the naming response of the target will be in the intersection of the sets generated by the prime and the target. When prime and target are unassociated, only the set generated by the target can contain the target but not the set generated by the prime. Therefore, the intersection of sets would not contain the target for unrelated prime-target pairs. This model by M. S. Humphreys and colleagues (1993) can explain evidence for a parallel activation of prime and target as well as the results of the current task.

A related theory, the retrieval account by Whittlesea and Jacoby (1990) seems also suited, in principle, to explain our results. Resembling the compound-cue theories, the essential new point of the retrieval account is that the prime is only utilized if this is of functional value for target processing. Accordingly, especially for degraded targets (e.g., pLAnT) of related prime-target pairs (e.g., if *green* is the prime), the prime and fragments of target processing are incorporated into a compound-cue to retrieve the correct target representation. Therefore, the retrieval account can explain the current findings straightforwardly, particularly because we used masked presentation. To reconstruct what was only briefly presented, the compound of the initial processing results of prime and target (i.e., fragments of prime and target) is a potent retrieval cue for both prime and target in the case of a related prime-target pair, but of no help in the case of an unrelated pair.

Both the theories by M. S. Humphreys and colleagues (1993) and Whittlesea and Jacoby (1990) do not need the assumption of mutual facilitation. It remains unclear, however, whether these theories can explain the evidence for parallel activation and mutual facilitation obtained by Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008; see *Introduction*).

In search of priming models that can account for their data, Schmitz and Wentura (2012) suggested linking working memory models (where two and more concepts can be held active at the same time) and semantic priming models. The present data reemphasize this suggestion. In Experiment 1b, a clear memory component is incorporated into the design. Defining the target after its onset, our task resembles the partial report technique introduced by Sperling (1960) in his seminal work on iconic memory. The partial report technique allows for estimating the span of a buffer store (i.e., iconic memory in Sperling's work, short-term/working memory in our case) with a single probe. That is, stimuli are encoded and only after a maintenance period does a backward cue indicate the to-be-identified item, which then has to be retrieved from memory. Thus, a legitimate description of the results of Experiment 1b is to assume a larger span in the case of related pairs compared to unrelated pairs.

The priming effect in Experiment 1b may indeed arise from the (working) memory component included in the task. If semantically similar concepts in working memory mutually facilitate each other, our working memory system might be the place where some semantic priming effects arise. This is in line with the assumptions by Schmitz and Wentura (2012) and the neurocomputational model by Davelaar and colleagues (2006).

The memory model by Davelaar and colleagues (Davelaar et al., 2005; Davelaar et al., 2006 see also Haarmann & Usher, 2001; Usher & Cohen, 1999) already accounts for mutual facilitation in working memory, albeit with a localist representation of concepts (i.e., a single node symbolizes a concept, comparable to semantic network theories). Basically, they assume that for related words a spread of activation can cause mutual facilitation, what is not the case for unrelated word-pairs. Davelaar and colleagues (2006) obtained evidence for mutual facilitation in working memory in two ways. They observed a better recall for pairs of related but not unrelated word pairs across serial positions in a list. In their model, it is assumed that the first word facilitates the activation of the second word and the second word facilitates the activation of the first word. Because the activation of the first word is facilitated even after its offset, the first word shows better recall than it would be expected on the basis of its serial position. This mechanism can explain a zigzag pattern that is observed for the memory of lists containing related but not unrelated word-pairs. Therefore, compelling evidence for a mutual facilitation of related concepts in working memory is provided in two ways: by a higher overall performance for related word pairs and the zigzag effect which is only observed in the related but not in the unrelated condition. Although this approach is closely related to priming research and theories, Davelaar and colleagues (2006) did not yet apply their model to semantic priming data; it might be worthwhile to do so. As already indicated earlier, another starting point to account for mutual facilitation effects might be an expansion of parallel-distributed working memory models (e.g., Raffone & Van Leeuwen, 2001, 2003; Raffone & Wolters, 2001; Vogel et al., 2001; Wolters & Raffone, 2008).

2.4.6 Future research

For future research on parallel activation and mutual facilitation, it might be worthwhile to take a two-pronged approach. First, it seems to be a legitimate claim that priming theories should not be contradictory to the assumptions of parallel activation and mutual facilitation. Therefore, evidence for these processes in priming can further contribute to establishing this assumption as a benchmark for the test of semantic priming theories. Additionally, the nature of the processes that constitute the mutual facilitation in priming can be addressed by future priming research. Second, besides priming research, working memory research could benefit from an investigation of automatic effects of (semantic) similarity providing further insight into the architecture of memory.

For priming research, we see several starting points for initiating new research. As pointed out earlier, focusing on the combination of priming and post-cue paradigms, a precise prediction can be made if a spread of activation as assumed by Collins and Loftus (1975) or J. R. Anderson (1976, 1983, 1993) underlies the mutual facilitation of related concepts. The more the onset of the target-cue is delayed, the higher should be the increase in the size of the relatedness effect (unless a potential maximum is reached). For longer delays, the received activation at the target node should add up resulting in a stronger effect of mutual facilitation. Additionally, it could be varied whether prime-target pairs are symmetrically associated or whether the association is primarily from the prime to the target. For symmetrically associated prime-target pairs, there should be a more rapid lexical-semantic activation-buildup of the target due to a stronger mutual facilitation than for asymmetrically associated prime-target pairs. This manipulation would affect a mutual facilitation that can take place in working memory (Davelaar et al., 2006). Potentially, it could also impact the encoding into long-term memory. This can be tested by delayed recall after a distractor task.

Similar to a variation of the onset time of the target-cue, studies varying the prime-target SOA in a post-cue task could further test the assumptions which can be derived from spreading activation theories. Another line of research can arise from the studies conducted by Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al. 2014; Wentura & Frings, 2008). They provide evidence for mutual facilitation and parallel activation of *evaluatively* congruent concepts in priming. However, their three-process model that explains their findings can predict mutual facilitation and parallel activation also for other variants of semantic relatedness. Nevertheless, in Experiment 4a and b of the study by Schmitz and Wentura (2012), no evidence for automatic mutual facilitation of concepts sharing a nonevaluative category (e.g., person or animal) was observed. Taking the current study and the research by Davelaar and colleagues (Davelaar et al., 2005; Davelaar et al., 2006 see also Haarmann & Usher, 2001; Usher & Cohen, 1999) into account, it is unlikely that the effect investigated by Schmitz and Wentura (2012) is valence-specific. Therefore, the design by Schmitz and Wentura could be used with stimuli from more distinctive semantic categories with a higher similarity of concepts within the category (e.g., foods and dogs) to provide evidence for a mutual facilitation of concepts sharing a non-evaluative category.

For working memory research, future research could address the question whether mutual facilitation can also be observed when chunking is absent or reduced. Besides using associated words, for which often a common category retrieval cue can be generated, future research could use evaluative congruency to introduce similarity, which led to mutual facilitation in the priming tasks by Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al. 2014; Wentura & Frings, 2008). Additionally, and similar to the studies by Wentura and colleagues, different dimensions of a single stimulus could be used to introduce similarity and to provide the task-relevant features (e.g., the use of faces as stimuli would allow to use the valence of a face as the similarity defining dimension but testing only the memory for the identity of a person).

In these two first experiments, we aimed to raise awareness for the importance of explaining semantic priming effects in a model that includes the parallel activation of prime and target concepts. Parallel distributed processing models might require modification to incorporate both mechanisms of parallel activation and mutual facilitation. Incorporating parallel activation and mutual facilitation into existing theories of semantic priming, and regarding these two phenomena as benchmarks that priming theories should be able to account for, is arguably a worthwhile endeavor. Furthermore, priming theories and working memory models might mutually benefit from integration because parallel activation of several concepts is highly plausible in both theoretical arenas. Thus, gaining further support for the hypothesis of mutual facilitation in working memory is a logical next step. Assuming that mutual facilitation in working memory is based on the same principles as the mutual facilitation of concurrently active related concepts in priming, mutual facilitation in working memory should also occur if relatedness is implemented by evaluative congruency. This question will be addressed in Experiment 2a-d, which is reported next.

3 Effects of evaluative homogeneity in working memory

3.1 Using evaluative faces

The following experiments were all designed to test whether there are beneficial effects of *evaluative* congruency in working memory. As outlined at the beginning, there are empirical as well as theoretical arguments to hypothesize a performance benefit for similar items. Experiment 1b provided evidence for beneficial effects of *semantic* similarity in a priming task in which a working memory component is incorporated. This effect is in line with the assumption of mutual facilitation processes in working memory, which can be predicted by theories incorporating spreading activation. Alternatively, the effect can be explained by some compound-cue theories. For working memory, the dual-store neurocomputation model (Davelaar et al., 2005, 2006; Haarmann & Usher, 2001; Usher & Cohen, 1999), which incorporates a spreading activation process, proposes mutual facilitation for semantically related stimuli. But why is the focus now on effects of *evaluative* congruency, which is a highly specific kind of semantic similarity?

There are several main reasons for investigating *evaluative* congruency. First, there is already evidence for parallel activation and mutual facilitation of *evaluatively* congruent concepts in priming (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008) and there is already fairly conclusive evidence for mutual facilitation of *semantically* related concepts in working memory (e.g. Davelaar et al., 2006). Thus, the missing piece is convincing evidence in line with the assumption that evaluative congruency leads to a mutual facilitation in working memory. This effect should be observed because theories assuming a spread of activation can be applied to both evaluative congruency as well as other forms of semantic overlap. Additionally, other theories used to explain semantic priming are usually also applied to explain evaluative priming (when it is stimulus-stimulus based evaluative priming, see e.g. Schmitz & Wentura, 2012). The second reason is that for strong semantically related material a common category retrieval cue and sophisticated guessing strategies could be used (Crowder, 1979; Poirier & Saint-Aubin, 1995). However, in the case of the more subtle evaluative congruency, these strategic influences become less likely. A third advantage of using evaluative congruency arises when face stimuli are used. In this case the identity of the face can be used as task-relevant, for example in a change detection paradigm (e.g., Jackson, Wu, Linden, & Raymond, 2009; Jackson et al., 2014; Jackson, Wolf, Johnston, Raymond, & Linden, 2008) and the emotion could be used as the similarity defining feature. This is an advantage due to several reasons. On the one hand, this again makes strategic influences on performance less likely, which are based on the conscious perception of similarity. Using the similarity to generate a common category retrieval cue (like saying to yourself that there were four *angry* faces in a congruent trial) would not lead to any performance boost. On the other hand, as far as we know, there is no prior research that has investigated the effects of any kind of similarity on working memory performance in which the similarity defining feature and the task-relevant feature were distinguished. Thus, our research can provide a new procedure that has arguably the central benefit that influences of strategic processes become less likely. Please note as well, that in all experiments that obtained a beneficial effect of perceptual similarity on working memory performance, the similarity defining feature was always task-relevant (Jiang, Lee, et al., 2016; Lin & Luck, 2009; Sims et al., 2012). Further, when it is possible to establish stable effects of congruency using change detection and face stimuli, the use of faces allows for a more precise investigation of the origin of potential congruency effects in subsequent research. For example, as next steps the emotional expression could be omitted during study or during test to analyze at which stage effects of relatedness arise without changing the task-relevant feature, that is, the identity of the faces. In addition, it can be assumed that processing of valence is prevented or reduced for inverted faces (Fox et al., 2000; McKelvie, 1995; Prkachin, 2003). Therefore, these stimuli can be used to control for influences of perceptual similarity because the perceptual similarity but not the overlap of the evaluative component of the stimuli stays the same when faces are inverted. A further reason for using valence as the similarity defining feature and specifically faces as stimuli can be derived from existing paradigms and findings by Jackson and colleagues (Jackson et al., 2009, 2014) that will be described in detail in the following paragraph.

In a series of experiments by Jackson and colleagues (2009), participants had better recognition performance in a visual working memory task when they had to remember faces expressing anger instead of faces showing a neutral expression, an effect they labeled *angry face benefit*. In detail, participants were presented (for 2 s) with arrays of between one and four faces of either angry, happy, or neutral expression. Following a one-second retention interval, a probe face was presented, and participants had to make a recognition judgment (i.e., whether this face was included in the learning array or not). Participants performed significantly better when faces at learning showed anger instead of a neutral (or happy) expression. The observed angry face benefit shows that the task-irrelevant emotion in their design (participants only had to remember the identities) had an influence on performance. Therefore, the change detection task with emotional faces seems to be a suitable candidate to observe effects of evaluative processing, and maybe evaluative congruency, even if the emotion is not task-relevant. Even more interesting is the fact that, except for one condition in Experiment 1, more than a single face was presented in the memory display. Of note, in this single face condition, the angry face benefit was missing. Thus, since displays including more than one face were always homogeneous with regard to expression, their studies provide evidence that emotional faces (more precisely angry faces) presented in the *context* of other evaluatively congruent faces are processed or maintained more efficiently than, for example, neutral faces in the context of other neutral faces.

The strongest angry face benefit seems to occur when neutral faces serve as control stimuli. Therefore the effects could be partially based on evaluative overlap: In the condition with several neutral faces, there is no evaluative overlap, because the neutral stimuli do not possess evaluative associations (Fazio, 2007). In trials with several angry faces, these stimuli share their evaluative component and mutual facilitation could occur. Interestingly, in a study highly similar to the one by Jackson and colleagues (2009), which was conducted by Langeslag, Morgan, Jackson, Linden and Van Strien (2009), memory performance was higher for angry *and* happy faces over neutral faces. However, this better performance for emotional faces was only observed at load three (i.e., three faces had to be memorized) but not at load one. Therefore, their study provides evidence that angry and also happy faces presented in the context of other evaluatively congruent faces are processed or maintained more efficiently.

Thus, the focal hypothesis tested in the following experiments is that evaluative overlap of items in a visual short-term memory task increases recognition performance. We should hasten to add that we do not claim that there is no pure angry face benefit. We acknowledge that there are studies with happy faces in the control condition (e.g., Experiment 1 by Jackson et al., 2009) that show an angry face benefit. Beyond that, later studies (Thomas, Jackson, & Raymond, 2014) found an angry face benefit even with a single memory item. Nevertheless, some of the reported effects may be due not solely to an angry face benefit, but also to a memory benefit for faces presented in an evaluatively congruent context. The following experiments in this thesis were also designed to replicate the angry face benefit. Nevertheless, with regard to the focus of this thesis, aiming for evidence in line with the assumption of a mutual facilitation in the case of evaluative congruency is more relevant.

3.2 Theoretical basis

There are several theories that suggest that several faces showing the *same* emotion should lead to enhanced working memory performance. These are the three process model by Schmitz and Wentura (2012), spreading activation theories and partially compound-cue models, which are described in more detail in the theory section of this thesis. In the three process model, it is assumed that evaluatively congruent concepts mutually facilitate each other's activation when they are active in parallel. The authors further suggest that this mutual facilitation process could operate in working memory. Further, the dual-store neurocomputational model (Davelaar et al., 2005, 2006; Haarmann & Usher, 2001; Usher & Cohen, 1999) proposes mutual facilitation processes for related concepts in working memory based on a spread of activation. Accordingly, activation would spread back and forth between evaluatively congruent stimuli (e.g., the representations of evaluatively valenced faces, sharing their emotion) so that they would be automatically maintained. For evaluatively incongruent stimuli, there would be no spread of activation back and forth between their representations. Therefore, they would be more likely forgotten.

3.3 Experiment 2a-d: Using change detection to investigate effects of evaluative homogeneity in working memory

We conducted a series of four experiments employing a change detection task using emotional (i.e., angry and happy) faces as stimuli to investigate the effects of evaluative congruency on working memory performance and, as a secondary aim, to replicate the angry face benefit (i.e., better performance for angry compared to happy and neutral faces) found in several studies (Jackson et al., 2008, 2009, 2014, but see Langeslag, Morgan, Jackson, Linden, & Van Strien, 2009). As we argued above, assuming that mutual facilitation can take place in working memory for parallel activated concepts, participants should perform better in evaluatively congruent compared to incongruent trials, as evaluative congruency can be assumed to be a specific kind of semantic relatedness. Furthermore, we aimed at investigating the effect of mere perceptual overlap by also applying conditions with inverted faces. For inverted faces, it is assumed that processing of valence is prevented or reduced (Fox et al., 2000; McKelvie, 1995; Prkachin, 2003). Therefore, by utilizing an inverted faces condition, we can control for effects of perceptual overlap. In addition, research on the angry face benefit can be enriched by validating whether this effect is observed in the control condition using inverted faces, which would indicate that the effect is based on perceptual overlap of the face stimuli.

3.3.1 General paradigm and procedure: Experiment 2a-d – Upright faces

Based on our theoretical assumptions, we adapted the paradigm of Jackson and colleagues (2009) in the following way: In each trial, four faces were shown for a brief period (2 s). We manipulated the emotional expressions of the faces in the following way. In congruent trials, all four faces showed either an angry or happy expression. In incongruent trials, two faces showed an angry expression and two a happy expression. After a short retention interval (1 s), the display was repeated (a) without any change or (b) with one face replaced by a new one of the same emotion (see Figure 4).[13] Participants had to indicate whether a change had taken place or not. This design allowed, first of all, to test for an emotional congruency effect: Is the performance higher in congruent compared to incongruent displays? Second, we can test for the angry face benefit found by Jackson and colleagues (2009) by comparing the performance in congruent angry trials with the performance in congruent happy displays.

We conducted four experiments with a total of N = 207 participants. Since differences between experiments were minor (see *Methods*), we report the four experiments in the format of a single study with (a) overall analyses employing experiment as a factor and, alternatively, (b) meta-analytic procedures to answer the question whether a congruency effect exists or not.

The main reason for this synoptical report is given by the chronology and outcome of the four experiments. In the first experiment, we observed an effect of d_Z = .58 for the impact of evaluative congruency on working memory performance, an effect of a medium size according to (J. Cohen, 1988). Therefore, Experiment 2b already

[13] Note, we deviated here from the procedure used by Jackson and colleagues (2009) who presented a single probe face. This procedure would have introduced a confound into our design because the emotion might serve as a retrieval cue. That is, in congruent trials, the angry test face, for example, must be compared with all four angry faces from the learning display. In incongruent trials, however, the angry face must only be compared with the two angry faces from the learning display and not with the happy faces. Therefore, a whole-probe recognition task was utilized. In change detection tasks, both versions can be found (e.g., Wheeler & Treisman, 2002 used both whole-probed and single probed recognition. In both classical studies by Luck and Vogel (1997) and Alvaraz & Cavanagh (2004) whole-probed recognition was used).

included an additional condition to explore a follow-up question: Besides the replication condition, we repeated the experiment by presenting inverted faces. Displays with inverted faces have the same perceptual features as upright faces; however, valence processing for inverted faces is reduced (McKelvie, 1995; Prkachin, 2003). Thus, if a beneficial effect is due to evaluative homogeneity and not due to perceptual overlap, congruency effects should be reduced or absent for inverted faces. In Experiment 2b, we added the factor upright versus inverted faces in a blocked, counterbalanced within-participants design. However, we observed carry-over effects in this experiment. The first blocks indicated the expected pattern of results, that is, numerically a congruency effect for upright faces (which, however, failed to be significant, potentially because of a reduced number of trials) and a null effect of inverted faces. The second blocks, however, mimicked the pattern of the first blocks although stimulus type (i.e., upright vs. inverted) was exchanged. Therefore, we realized the factor stimulus type as a between-participants factor in Experiment 2c. Obtaining a clear null result (for upright faces), we wondered about the existence of the congruency effect and conducted a fourth experiment. The congruency effect was positive but again not significant. However, to get to the point, a meta-analytic approach still provides evidence for a small effect of congruency for upright faces (and no effect for inverted faces).

For this report, we decided to report the full set of results in the following way. We will first report the results of the four samples testing for congruency effects of upright faces to answer the question whether there is such an effect or not. Because of the carry-over effect and to keep comparability to the other three experiments, we restricted Experiment 2b to the subsample that started with the upright faces block. Besides meta-analytic evidence for the congruency effect, we were able to replicate the angry face benefit effect.

Subsequently, we will report the results of the two samples that tested for the congruency effect and the angry face benefit effect of inverted faces (i.e., the subsample of Experiment 2b that started with the inverted block, labeled Experiment 2b.2, and the corresponding "inverted faces"-subsample of Experiment 2c, which is labeled in the following 2c.2). Finally, to give full transparency, we will give a detailed report of the carry-over effect observed in Experiment 2b in Appendix A.

3.3.2 Materials and methods

Participants. In Experiments 2a to 2d, 207 students (149 females, age range 18-35 years, *Md* = 24) participated and were paid 6 to 8 € for participation. All of them had normal or corrected-to-normal vision. The data of five further participants were excluded because they did not perform significantly above chance. Table 1 provides a detailed description of the demographic data for the four samples.[14]

Table 1 *Demographic data, sample of the main analysis for upright faces (Experiment 2a-d)*

	Experiment 2a	Experiment 2b	Experiment 2c	Experiment 2d
N	38	34	37	98
Gender (f/m)	29/9	25/9	22/15	73/25
Median age (Range)	24 (21-32)	24 (18-30)	23 (18-35)	23 (18-35)

Design. Essentially, evaluative congruency (congruent vs. incongruent) was manipulated within participants. Evaluative congruent displays were composed of either four angry or four happy faces; evaluative incongruent displays were composed of two angry and two happy faces. Furthermore, using a change detection task, change and no change trials were utilized: Whereas in half of the trials, the same faces were presented at encoding and test, in the other half of trials a single face was replaced at test by a different face expressing the same emotion.

[14] Because similar effects as in evaluative priming are expected, the priming effect that is observed when nonverbal primes are used (d_Z = .48) or the effect with nonverbal targets (d_Z = .36) might provide a first guess for an initial power calculation. When an effect size of d_Z = .42, which is between these two values, is assumed, 37 participants would be required to achieve a power of 1-β = .80.

Materials. Grayscale cut-out versions of faces from the Karolinska Directed Emotional Faces databank (KDEF) were used (Lundqvist, Flykt, & Öhman, 1998). Six (Exp. 2a) or 12 (Exp. 2b to 2d) male identities expressing happiness and anger were selected. In Experiment 2a, trials employing neutral expressions were added (see *Procedure*).

There were some further minor differences between experiments: In Experiment 2a-c, angry and happy faces presented to a given participant depicted the same persons, whereas in Experiment 2d different identities were used for angry and happy faces (counterbalanced across participants). Whereas in Experiment 2a only six stimulus persons were employed, in Experiment 2b to 2c two sets of six stimulus persons were used across participants (i.e., in Experiments 2b and 2c, a half of the participants received Set 1 or Set 2, respectively; for Experiment 2d, see above).

Procedure. Stimuli were presented on 17-inch monitors (1024x768 pixels) using E-Prime 2 software (Schneider, Eschman, & Zuccolotto, 2002) and were viewed from a distance of approximately 70 cm. Each picture subtended 2.82° × 3.95° visual angle. All four faces of the learning display as well as all four faces of the test display covered together a visual angle of 6.07° × 8.35°. In each trial four faces were presented both at encoding and at retrieval, applying a whole display recognition task (Rouder, Morey, Morey, & Cowan, 2011).

Figure 4 illustrates the trial procedure in the change detection task. Following the sequence "+", "×", "+", which should prepare participants, the four faces of the learning display were presented for 2000 ms and then replaced by a screen showing a fixation cross, presented for 1000 ms. The subsequent test display, also containing four faces, remained on the screen until a response was given. Participants indicated by keypress whether a single face identity was replaced from study to test or whether the same faces were presented again. During practice, feedback ("Error") was presented after false responses.

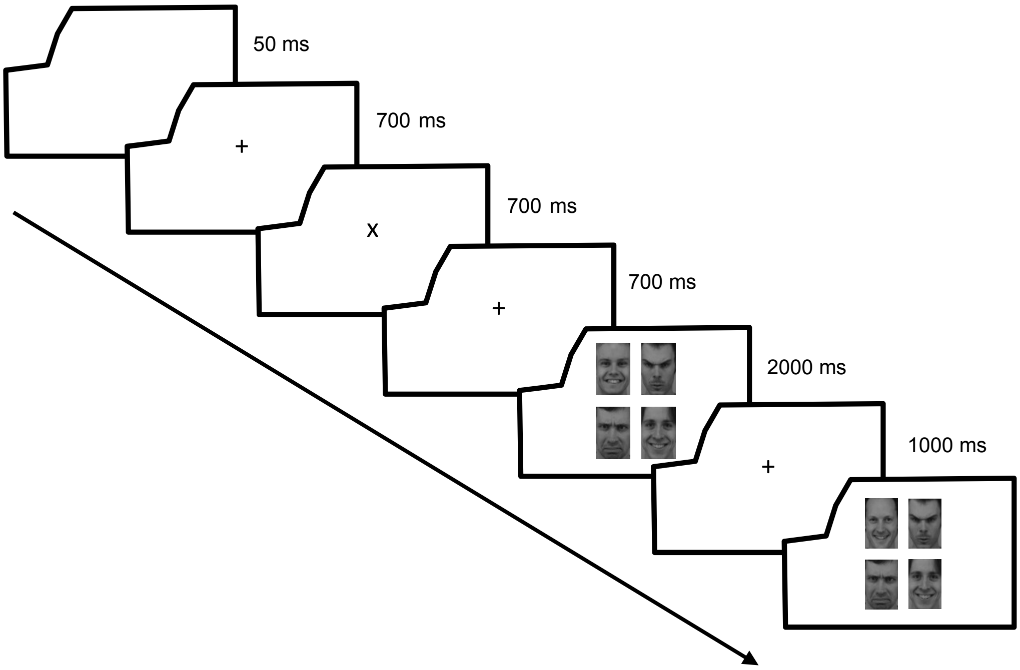


Figure 4. *Trial sequence for the Experiments 2a–2d. (Depicted is a change trial).*

Each participant was presented with six angry faces and six happy faces throughout the experiment (see *Materials*). Construction of the trial sequence was as follows: In a given trial, one of the twelve faces was the designated target. (To prevent a misunderstanding: the target status was not evident for the participants.) It was randomly assigned to one of the four locations. In congruent trials, the three distractor positions were filled by three expressions of the same emotion as the target, which were randomly selected from the respective set. Accordingly, in incongruent trials, one expression of the same emotion as the target and two expressions of the other emotion were randomly selected from the sets. In change trials, a further expression of the same emotion as the target was randomly selected from the respective set to replace the target in the test display. In both congruent and incongruent trials, in half of the trials the target was an angry face; in the other half of the trials the target was a happy face. A block of trials consisted of 24 trials (2 [congruent vs. incongruent] × 2 [change vs. no change] × 6 repetitions) plus 6 trials showing faces with a neural emotional expression in Experiment 2a. The identities used were drawn randomly in each trial.

The experiment started by a practice phase of one block in Experiment 2a which was followed by the main phase (240 trials [i.e., 10 blocks] plus 60 neutral trials in Experiment 2a with a short break after 5 blocks, 192 trials [i.e., 8 blocks] in Experiment 2b with a break after 4 blocks, 288 trials [i.e., 12 blocks] in Experiment 2c and 2d with breaks after 4 and 8 blocks). After each block, feedback was presented, showing the percentage of correct responses in this block. Furthermore, participants that did not reach a performance above 60 % were requested to try harder to remember the faces, reminding them that a performance of 50 % correct responses can be achieved by mere guessing. At the end of each experiment, participants filled in a short post-experimental questionnaire (measuring task difficulty, comprehensibility of instructions, etc.).

3.3.3 Results

Congruency effect. Table 2 shows the mean performance in the congruent and the incongruent condition as well as the difference between these two conditions – the average congruency effect – quantified in *d'*. (To account for relative hit or false-alarm rates of zero or one, the *log-linear* correction was chosen, see Hautus, 1995). As is evident from the 95%-confidence intervals for the congruency effects of all four experiments, only Experiment 2a showed a significant effect.

However, a mixed 2 × 4 ANOVA with congruency (congruent vs. incongruent) as a within-participants factor and experiment as a between-participants factor yielded the expected significant congruency effect, $F(1,203) = 6.64$, $p = .011$, $\eta_p^2 = .032$, providing evidence for better performance in the congruent compared to the incongruent condition. This effect was not moderated by the factor experiment, $F(3,203) = 2.03$, $p = .110$, $\eta_p^2 = .029$ ($F(3, 203) = 2.20$, $p = .089$, $\eta_p^2 = .031$ for the main effect of experiment). Since in this analysis the experiments are all weighted equally irrespective of the number of participants, we tested the congruency effect in the overall sample, disregarding the factor experiment; it remained significant, $t(206) = 2.40$, $p = .017$, $d_Z = .17$.

Table 2 *Mean performance (in d') for upright faces as a function of congruency in the Experiments 2a-d.*

Exp	*N*	Congruency effect[a] [95% CI]	Congruent Mean *d'* (*SD*)	Incongruent Mean *d'* (*SD*)
2a	38	0.16 [0.07 – 0.26]	1.44 (.48)	1.28 (.47)
2b	34	0.13 [-0.03 – 0.29]	1.67 (.63)	1.54 (.61)
2c	37	-0.03 [-0.16 – 0.09]	1.55 (.53)	1.58 (.58)
2d	98	0.04 [-0.04 – 0.12]	1.44 (.52)	1.40 (.54)
Total	207	0.06 [0.01 – 0.12]	1.50 (.54)	1.43 (.55)

[a] Mean congruency effect = *d'*(congruent) – *d'*(incongruent)

As a third way of analyzing the data, we provide a forest plot (Cumming, 2012) together with a meta-analytical investigation of the overall effect (see Figure 5). We based the analysis on the means and the standard deviations of the congruency effect (i.e., the performance in the congruent condition minus the performance in the incongruent condition). Analyses were performed in ESCI (Cumming, 2012). Due to the close comparability between studies, the fixed effect model was used (Cumming, 2012). (With $T = 0.07$ and a 95% CI of [0, 0.14] for τ, results were rather homogeneous, and the test for the null hypothesis of homogeneity did not provide a significant result, $p > .05$.) The overall congruency effect (see Figure 5) given by this analysis, $M_{\Delta} = 0.08$ was significantly different from zero, $t(206) = 3.00$, $p = .003$, $d_Z = .21$.

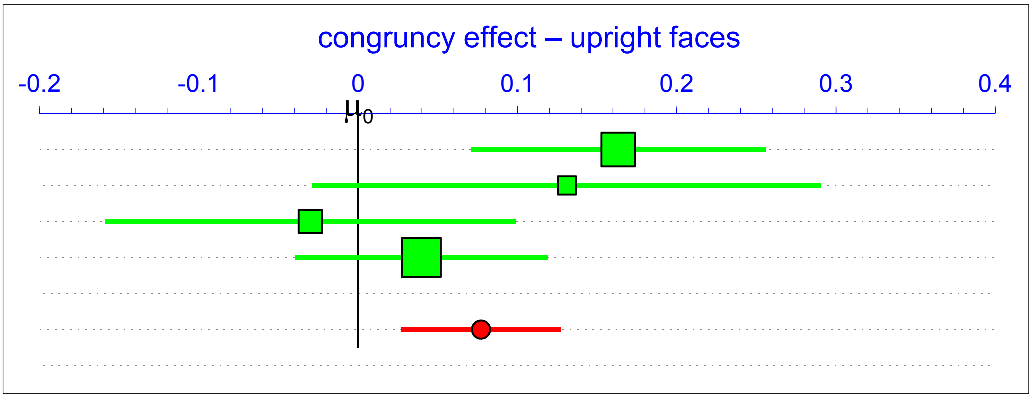


Figure 5. *Forrest plot of the average congruency effect for upright faces (Experiment 2a-d). From the top to the bottom, confidence intervals of the Experiments 2a–2d and of the overall congruency effect are plotted. The congruency effect was the difference between the d' values of congruent and incongruent trials. The size of the squares indicates the weight of the individual experiments in the analysis.*

Angry face benefit. We also conducted analyses with the aim of replicating the angry face benefit found by Jackson and colleagues (2009). Table 3 provides the overview of the effects for the individual experiments. We performed a 2 (emotion) × 4 (experiment) mixed ANOVA comparing the performance in trials with only angry faces versus trials with only happy faces.[15] This analysis revealed a significant effect of the factor emotion, $F(1,203) = 8.60$, $p = .004$, $\eta_p^2 = .041$. This effect was not moderated by the factor experiment, $F(3,203) = 0.56$, $p = .644$, $\eta_p^2 = .008$ ($F(3,203) = 1.86$, $p = .138$, $\eta_p^2 = .027$ for the main effect of experiment). Moreover, this effect was further confirmed by an analysis without the factor experiment, comparing the performance for trials with only angry and only happy faces, $t(206) = 3.26$, $p = .001$, $d_Z = .23$.

[15] As for the analysis of the congruency effect, we restricted Experiment 2b to the subsample that started with the upright faces block to keep comparability to the other three experiments.

Table 3 *Mean performance (in d') for upright faces as a function of face valence (angry face benefit) in the Experiments 2a-d*

Exp	*N*	Angry face benefit[a] [95% CI]	Angry faces Mean *d'* [b] (*SD*)	Happy faces Mean *d'* [c] (*SD*)
2a	38	0.24 [0.07 – 0.41]	1.56 (.59)	1.32 (.32)
2b	34	0.09 [-.17 –.34]	1.72 (.70)	1.63 (.76)
2c	37	0.08 [-.13 –.29]	1.59 (.61)	1.51 (.63)
2d	98	0.14 [.02 – .25]	1.51 (.57)	1.37 (.63)
Total	207	0.14 [.05 – .22]	1.57 (.60)	1.43 (.63)

[a] Mean angry face benefit = *d'*(only angry faces) – *d'*(only happy faces)

[b] performance in angry only trials

[c] performance in happy only trials

A meta-analytical approach confirmed the result (see Figure 6). Basing the analysis on a fixed effect model, there was no evidence for heterogeneity in the data (τ = 0, CI = [0, 0.175972], p = .5823 for testing the null-hypothesis of homogeneity). In the analysis, a significant angry face superiority effect of M_{Δ} = 0.15 was observed, $t(206)$ = 3.59, $p < .001$, $d_Z = .25$.

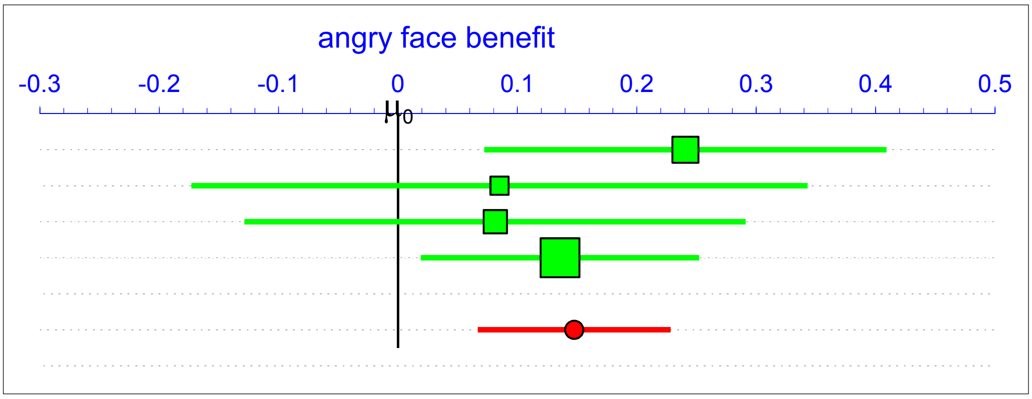


Figure 6. *Forrest plot of average angry face benefit for upright faces (Experiment 2a-d). From the top to the bottom, confidence intervals of the Experiments 2a–2d and of the overall angry face benefit are plotted. The angry face benefit was calculated as the difference between the d' values of trials with only angry faces and trials with only happy faces. The size of the squares indicates the weight of the individual experiments in the analysis.*

For the sake of completeness, trials with neutral faces (which were only used in Experiment 2a) were associated with a mean performance (in *d'*) of *M* = 1.12 (*SD* = 0.54). This performance significantly deviated from the performance for happy faces (see

Table *3*), $t(37) = 3.22$, $p = .003$, $d_Z = .52$, indicating that, in our paradigm, happy faces are also processed more efficiently than neutral faces.

3.3.4 Interim discussion

The overall analysis of the four experiments gives modest support for our hypothesis. We found significantly enhanced performance due to evaluative congruency in a standard working memory task. The finding was not moderated by the factor experiment, indicating a reliable effect that is not influenced by minor procedural changes. Hence, evaluative congruency seems to have the same influence on working memory performance as it has on performance in typical priming tasks: it leads to a more efficient processing. However, admittedly, the effect is only of small size.

A further result is the replication of the angry face benefit. Overall, evidence for this benefit is stronger than the evidence for the congruency effect. However, in our experiment, the effect was rather small as well.

As already indicated in the overview, a follow-up question is whether the congruence effect and the angry face benefit can be interpreted as caused by the emotional content of the faces or by perceptual features. One means to provide evidence for the former or the latter is to explore whether the same effects found with upright faces will be found with inverted faces as well. If this happens, one might argue (with some caution, see, e.g., Horstmann & Bauland, 2006 for a critical argument) that the visual features that constitute a specific emotional expression are primarily responsible for the effect.

3.4 Experiments 2b.2 and 2c.2: Inverted faces

3.4.1 Materials and methods

A total of 70 participants (34 in Experiment 2b.2, 36 in Experiment 2c.2, 46 female, age range 18 – 35, *Md* = 23) participated in Experiments 2b.2 and 2c.2. The data of two further participants were excluded because they did not perform above chance. Design, materials, and procedure were exactly the same as in Experiments 2b and 2c, except that all faces were presented inverted.[16]

3.4.2 Results

Congruency effect. Table 4 shows the descriptive statistics for the experiments presenting inverted faces.

Table 4 *Mean performance (in d') for inverted faces as a function of congruency in the Experiments 2b.2–2c.2*

Exp	*N*	Congruency effect[a] [95% CI]	Congruent Mean *d'* (*SD*)	Incongruent Mean *d'* (*SD*)
2b.2	34	0.00 [-0.16 – 0.16]	1.07 (.43)	1.07 (.54)
2c.2	36	0.08 [-0.04 – 0.20]	1.13 (.55)	1.05 (.47)
Total	70	0.04 [-0.05 – 0.13]	1.10 (.49)	1.06 (.50)

[a] Mean congruency effect = *d'*(congruent) – *d'*(incongruent)

In a 2 (congruency) × 2 (experiment) mixed ANOVA there was no congruency effect, $F(1, 68) = 0.71$, $p = .401$, $\eta_p^2 = .010$, no main effect for experiment, $F(1, 68) = .03$, $p = .87$, $\eta_p^2 < .00$, and no interaction $F(1, 68) = .73$, $p = .397$, $\eta_p^2 = .011$.

[16] Because of a carry-over effect in the analysis of the congruency effect, we restricted Experiment 2b.2 to the subsample that started with the inverted faces block. For the purpose of consistency, the analysis of the angry face benefit for inverted faces also only included the data from participants that started with an inverted faces block.

For Experiment 2b and 2c, the direct comparison of upright versus inverted faces in a 2 (stimulus type: upright vs. inverted) × 2 (congruency) × 2 (experiment) mixed ANOVA yielded no significant interaction of stimulus type and congruency, $F(1, 137) = 0.02$, $p = .884$, $\eta_p^2 < .000$ that is, we cannot claim that the congruency effect for inverted faces is significantly smaller than the one for upright faces.

Angry face benefit. Table 5 shows the descriptive statistics for the angry face benefit if inverted faces were presented.

Table 5 *Mean performance (in d') for inverted faces as a function of face valence (angry face benefit) in the Experiments 2b.2–2c.2*

Exp	*N*	Angry face benefit[a] [95% CI]	Angry faces Mean *d'* [b] (*SD*)	Happy faces Mean *d'* [c] (*SD*)
2b.2	34	0.46 [.21 – .71]	1.30 (.57)	.84 (.55)
2c.2	36	0.48 [.28 – .68]	1.37 (.60)	.89 (.65)
Total	70	.47 [.31 – .63]	1.33 (.58)	.86 (.60)

[a] Mean angry face benefit = *d'*(only angry faces) – *d'*(only happy faces)

[b] performance in angry only trials

[c] performance in happy only trials

A 2 (emotion) × 2 (experiment) mixed ANOVA revealed better performance for trials with only angry compared to only happy inverted faces, $F(1,68) = 35.74$, $p < .001$, $\eta_p^2 = .345$. This effect was not moderated by the factor experiment, $F(1,68) = .01$, $p = .919$, $\eta_p^2 < .001$ ($F(1,68) = .24$, $p = .623$, $\eta_p^2 = .004$, for the main effect of experiment).

The direct comparison of upright versus inverted faces in a 2 (stimulus type: upright vs. inverted) × 2 (emotion) × 2 (experiment) mixed ANOVA yielded a significant interaction of stimulus type and emotion, $F(1,137) = 11.69$, $p = .001$, $\eta_p^2 = .079$. That is, the angry face benefit for inverted faces is even larger than the one for upright faces.

3.4.3 Interim discussion

Analyzing the congruency effect with inverted faces, no significant influence on performance was observed. Displays with inverted faces, which have the same perceptual features as upright faces but reduced valence (McKelvie, 1995; Prkachin, 2003), lack a strong evaluative congruency that might cause the effect for upright faces. Although tested with a rather small sample, the absent effect for inverted faces can provide a first hint about the underlying processes. Seemingly, the effect for upright faces (if taken for granted) is at least not merely caused by perceptual overlap. Of course, we should hasten to add that any conclusions must be made with caution since there are overall no differences between the congruence effects for upright faces and inverted faces in the two experiments that included the factor stimulus type (upright vs. inverted).

Interestingly, the angry face benefit for inverted faces was significant, and it was even larger than the one for upright faces. We will return to this issue in the discussion below.

3.5 Discussion

The experiment yielded two important results. First, we found evidence for an evaluative congruence effect in a visual working memory task. If the memory display was homogeneous with regard to the facial expressions, change detection performance was higher compared to heterogeneous displays. Admittedly, the effect is small and only significant in an analysis that combined all four experiments. Second, we conceptually replicated the angry face benefit. Finding this benefit with inverted faces as well will enrich the discussion about this effect. The two results will be discussed successively.

3.5.1 Congruence in working memory

The most sensible explanation of the observed congruency effect might be mutual facilitation of evaluatively congruent concepts. Spreading activation theories (J. R. Anderson, 1976, 1983, 1993; Bower, 1991; Collins & Loftus, 1975) assume that activation from an activated concept spreads to other related concepts, which are therefore preactivated. It can be assumed that once two related concepts are active at the same time, activation spreads back and forth between related (or evaluatively congruent) concepts. Hence, these two concepts would mutually and automatically maintain each other's activation. Furthermore, interpreting working memory as the activated part of long-term memory (Cowan, 2001; Oberauer, 2002, 2006, 2009a; Oberauer & Lange, 2009; Oberauer et al., 2013) mutual facilitation could be considered a working memory phenomenon. Indeed, mutual facilitation of related concepts was proposed to occur in a limited capacity short-term buffer (Davelaar et al., 2005, 2006; Haarmann & Usher, 2001; Usher & Cohen, 1999). In addition, for evaluatively congruent concepts, mutual facilitation was proposed by Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014). The authors provide evidence for mutual facilitation using priming and flanker tasks. Importantly, they suggest that this process, observed in priming paradigms, might take place in working memory.

Besides mutual facilitation due to a spread of activation, compound-cue theories could also arguably account for the current finding when it is not assumed that effects of evaluative overlap, in general, cannot be explained by compound-cue accounts (Schmitz, 2012). However, compound-cue theories might have problems to explain priming when there are no direct associations between stimuli (Lucas, 2000). Be that as it may, the compound-cue theory by Ratcliff & McKoon (1988) suggests that relatedness leads to a higher familiarity signal for compounds. When there is a higher familiarity signal for old faces learned in a congruent trial compared to an incongruent trial, the stimuli might have been recognized better in congruent compared to incongruent trials. That is because a correct no change response in our change detection task can be based on the impression that all stimuli of the test display have a strong familiarity signal. When one stimulus has a reduced familiarity signal, a change response could be given. Therefore, correct responses could also be based merely on the familiarity signal. Accordingly, a boost of the familiarity signal for the learned stimuli in congruent trials (but not incongruent trials) could lead to a higher performance. Please note, however, that the theory by Ratcliff & McKoon (1988) in its most common implementation has problems to account for the data obtained in Experiment 1a and 1b, in which evidence in line with mutual facilitation and parallel activation in the case of semantic relatedness in priming was obtained using a perceptual identification task. Deriving predictions for the current data from the theory by M. S. Humphreys (1993) would be even more speculative. While the generation of associative sets is plausible for words (in this case, the associative sets contain associates of the words) it is unclear whether such a set can be formed for face stimuli and what this associative set of a single face would contain if it existed. Therefore, the theory seems to not have the potential to predict the current findings. A remaining candidate for an alternative explanation different to mutual facilitation based on spreading activation processes could be the retrieval account by Whittlesea and Jacoby (1990). For this theory again, it is unclear whether it can account for the kind of similarity, namely evaluative congruency, which was manipulated in the current study. Nevertheless, a potential application of this theory is considered in more detail in the discussion of Experiment 4. To summarize, a successful explanation of the current findings by current compound-cue theories remains a yet unresolved riddle.

Instead, the simplest account for the current evidence seems to be mutual facilitation in the case of relatedness that can be predicted by spreading activation processes. There are several caveats, though. First, in contrast to the mutual facilitation assumption, one might hypothesize that a homogeneous memory set might produce a kind of "sharpening" (i.e., an increase in precision) of memory representations. For example, Lin and Luck (2009), who found better visual working memory performance for color patches that were highly similar in hue compared to highly dissimilar color patches (see section 1.2.2), argued that lateral inhibition within color space might possibly have caused this sharpening. Correspondingly, one might assume that these precision-enhancing mechanisms are at work as well in our face experiments. It is up to further research to disentangle between mutual facilitation and sharpening processes. The same holds true for other explanations suggested by Lin and Luck (2009).

The second caveat concerns the location of the effect. In the paragraphs above we favored an interpretation in terms of mutual facilitation (or sharpening) in the maintenance phase of a trial. Of course, it is not precluded by our design that facilitation processes occur at either encoding or retrieval. In this regard, however, Experiment 3 by Jackson and colleagues (2014) is informative. Participants were presented with either only angry or only happy faces at encoding and with a single neutral face at test. Participants indicated whether an old or a new identity was presented at test. Most importantly, during the maintenance phase a positive or negative word was presented. Interestingly, participants performed better in the change detection task when valence of the encoded faces and word valence matched. Hence, their experiment provides evidence for a beneficial effect of evaluative congruency in working memory that originates during maintenance.

The third caveat is given by the carry-over effect found in Experiment 2b (see Appendix A). In Experiment 2b, a (numerically) better performance in congruent compared to incongruent trials was only observed when participants first performed the task with upright faces. If participants firstly remembered inverted faces, they showed the reversed pattern. First of all, this carry-over effect fits the overall impression that the congruence effect is a fragile phenomenon. However, there is one speculative interpretation that deserves further investigation. Processing of upright faces is typically interpreted in terms of global holistic processing, whereas processing of inverted faces is seen as dominantly component-based (Bartlett & Searcy, 1993; Carbon & Leder, 2005; Rhodes, Brake, & Atkinson, 1993; Searcy & Bartlett, 1996; Wilford & Wells, 2010). In this regard, it might be the case that the type of the first block of trials determines the processing mode (i.e., holistically for upright faces, component-based for inverted faces) which then carries over to the second block. Indeed, it can be assumed that feature based or holistic processing can be experimentally induced. Therefore, inverted faces might have led to a less holistic processing, similarly as the local processing of Navon stimuli can disrupt global processing in other following tasks (Z. Gao, Flevaris, Robertson, & Bentin, 2011; Liberman & Förster, 2009). Please note, however, that it is an open question whether these inductions can last for longer periods. However, when this can be assumed, the carry-over effect in Experiment 2b (see Appendix A) can potentially be explained by a reduced holistic processing, when upright faces follow inverted faces. This assumption is however rather speculative and it is plausible to assume that such effects could be restricted to the evidence for effects of similarity in the visual domain. If the assumption holds true, in collectivistic cultures in which holistic thinking is rather predominant (Nisbett & Miyamoto, 2005), similarity should boost performance more than when westerners from an individualistic culture perform the same task.

A final caveat is concerned with the smallness of the effect. One feature that might contribute to this meagerness is that even in incongruent trials there is always one congruent item for each stimulus of the memory set (since there were always two happy and two angry faces). Although this was a necessity of experimental design – an emotion singleton (e.g., one happy face in a context of three angry faces) would have been outstanding in the memory set – the representation of each face might have been boosted a bit by the congruent face.

Another feature which might have led to the rather small effect size is that valence was not task-relevant. In this case, one might argue that strong effects of evaluative overlap, without any need to process the emotion, would have been rather remarkable. Following this argumentation, we tried to implement a procedure in Experiment 3a-c that has the potential to overcome the possible reductions of the effect size.

3.5.2 The angry face benefit

Another major point of the current series of experiments is the replication of the angry face benefit. This effect seems to be a robust phenomenon since it turned out even in a slightly different experimental setting. Interestingly, in our design, there was even a stronger angry face benefit effect for inverted faces. This finding differed from the result of Experiment 5 by Jackson and colleagues (2009) in which there was no angry face benefit if faces were inverted. It will be up to further research to clarify this discrepancy, which might be due to procedural alterations. For example, Jackson and colleagues (2009) used two and four face displays in their study. The difference in the angry face benefit between upright and inverted faces seems to hold especially for two-face displays but to a lesser extent for four-face displays (see Figure 4 in Jackson et al., 2009). It is important to note that even if it would turn out that the angry face benefit can be reliably found with inverted faces (as indicated by our results) this would not trivialize the effect. In a somewhat different context, Horstmann and Bauland (2006) convincingly argued that specific emotional expressions might have evolved as they did because the features of the expression are more easily detected or processed. In our present ecology, there is then a solid correlation between these easy-to-process features and threat because most of the time we see these features in an upright face that clearly conveys the threat. Thus, even if the angry face benefit in working memory might have evolved because it was advantageous to prioritize angry faces, the triggers for the beneficial memory processes might be the features of a typical angry face. If this is the case, the angry face benefit might hold even for inverted faces since the triggering features are processed in this case as well.

To summarize, with some caution we found evidence for a boost in working memory performance due to evaluative congruency. This effect is likely caused by the mutual facilitation of evaluatively congruent representations in working memory. Accordingly, working memory phenomena resulting from semantic relatedness or evaluative congruency and some priming effects are likely based on common mechanisms. Thus, a more general framework like the model by Usher and colleagues (Davelaar et al., 2005, 2006; Haarmann & Usher, 2001; Usher & Cohen, 1999) suggests itself for a broader application in both areas of research.

4 Effects of evaluative homogeneity in change detection with locational changes and reduced set size

Why was the effect observed in Experiment 2a-d rather small? As partially outlined above, there might be some minor caveats that could have, in sum, led to a reduced effect size. Possibly, these caveats can be overcome to obtain a robust benefit for evaluatively congruent items compared to evaluatively incongruent items in classic working memory tasks. Therefore, all following experiments were originally based on the same assumptions as the experiments before, namely mutual facilitation of evaluatively congruent concepts in working memory (Schmitz & Wentura, 2012; Schmitz, Wentura, & Brinkmann, 2014). But what precisely could have reduced the effect in Experiment 2a-d?

First of all, and as already mentioned, there was congruency in the incongruent condition in the Experiment 2a-d: In congruent trials, one face was accompanied by three stimuli sharing valence. Accordingly, in congruent trials, there should be a high quantity of mutual facilitation. In incongruent trials, there were two positively valenced faces and two negatively valenced faces. Thus, one face is accompanied by only one other stimulus of the same valence. Still, there is one stimulus that is evaluatively congruent. Accordingly, there is one stimulus potentially maintaining the other one and there should still also be mutual facilitation in incongruent trials. This is not necessarily a problem, because there should be more mutual facilitation in congruent compared to incongruent trials. However, the beneficial effect of mutual facilitation possibly has some limits. When the benefit by additional evaluatively congruent concepts is not linear and when it has an asymptote, three helping evaluatively congruent concepts might be not that much more helpful than only one evaluatively congruent concept. However, in contrast, one evaluatively congruent concept could be much more helpful than no evaluatively congruent concept. Accordingly, it is apparent that the same task as in Experiment 2a-c with set size two instead of four could lead to an enhanced effect of evaluative congruency. Thus, in all following experiments, variants of the change detection task with set size two were adopted.

The second concern over a possible reduction of a congruency effect can be derived from the feature-specific attention allocation framework by Spruyt and colleagues (Everaert, Spruyt, & De Houwer, 2011; Spruyt, De Houwer, & Hermans, 2009; Spruyt, De Houwer, Hermans, & Eelen, 2007, for evidence see also Gast, Werner, Heitmann, Spruyt, & Rothermund, 2014; but see M. Becker, Klauer, & Spruyt, 2016 for an ambiguously interpretable reexamination).[17] In Experiment 2a-d of the current thesis, valence was not task-relevant. Thus, perfect performance could have been achieved by remembering all identities and ignoring the facial expressions and the emotions of the faces completely. The feature-specific attention allocation framework, however, suggests that it is mandatory in evaluative priming to assign attention to the affective stimulus dimension in order to observe a significant evaluative priming effect (e.g., Spruyt et al., 2007). When participants had to respond to the stimulus valence in the majority of trials in a priming task, an evaluative priming effect emerged in trials implementing evaluative categorization as well as trials using a non-affective semantic categorization task. In contrast, the evaluative priming effect disappeared (even in the evaluative decision task) when participants assigned attention to non-affective stimulus features in 75% of trials (e.g. Spruyt et al., 2007). In our case, facial identity can be regarded as the non-affective stimulus feature. This is especially the case if the assumptions of the model of face processing by Bruce and Young are taken for granted, which assume that facial identity and emotion are processed independently (Bruce & Young, 1986, but see, e.g. Ellamil, Susskind, & Anderson, 2008; Fisher, Towler, & Eimer, 2016; Fitousi & Wenger, 2013; Lander & Butcher, 2015). Please note that there is good reason to assume that the valence of the faces in Experiment 2a-d was processed at least to a certain degree: The valence of the faces led to a (small) evaluative

[17] see Everaert, Spruyt and De Houwer (2016) for evidence from the conceptually similar affect misattribution procedure.

congruency effect and a significant angry face benefit, with the last effect also present in the similar task by Jackson and colleagues (2008, 2009, 2014).[18]

4.1 Experiment 3a-c: Using change detection with set size 2 and locational changes to investigate effects of evaluative homogeneity

How precisely were these issues addressed in Experiment 3a-c? First, as anticipated above, congruency in the incongruent condition was removed by using set size two. Thereby, the effect of congruency can potentially be boosted, given that the other task characteristics remain virtually the same as in Experiment 2a-d. Of course, we could not simply reduce the set size from four to two without changing other parameters as well because the performance of participants could otherwise be at ceiling. Changing locations of stimuli from study to test as well as changes in the arrangement lead to a drop in memory performance (e.g. Boduroglu & Shah, 2009; Jiang, Olson, & Chun, 2000; Logie, Brockmole, & Jaswal, 2011; Mutluturk & Boduroglu, 2014[19], but see Woodman, Vogel, & Luck, 2012). Therefore, to enhance task difficulty, in each of the Experiments 3a-c, we presented the faces in different locations in the test display compared with the study array. In Experiment 3a, participants additionally remembered two shapes before seeing the learning display with two faces. Thereby, we introduced a

---

[18] Alternatively, these effects were merely based on the overlap of perceptual features and on the difference of perceptual features between angry and happy faces. In any case, at least the emotional expressions were processed and caused effects. Nevertheless, the valence of the expression could have been ignored.

[19] Logie, Brockmole, and Jaswal (2011) observed no effect of task-irrelevant changes of locations on performance when long retention intervals were used. We choose seemingly comparable study-test intervals; however, with faces we used more complex stimuli for which processing can be assumed to take longer. In addition, pre-tests of our task also suggest that changing the location from study to test still led to higher task difficulty.

higher load to further enhance task difficulty, a change which was however dropped in Experiment 3b and 3c.

The second issue, the potentially missing assignment of attention to the valence of the stimuli was addressed by a newly introduced task-characteristic in Experiment 3a and 3b. In short, in one-third of the trials, the memory for the emotion of one of the two faces was tested. One might argue that testing the memory for the emotion of the faces in only one-third of the trials is not enough to achieve a focus on valence. As outlined above, Spruyt and colleagues (2007) needed 75 % of trials with a focus on valence to observe an evaluative priming effect. Therefore, in the following experiments, it was *not* indicated at the beginning of a trial whether the emotion or the identity of a face will be tested. Only at test, participants knew which task they should perform. A central advantage of this implementation is that participants have to attend to the identity *and* the emotion of both faces in *every single trial*. This strong focus on valence was not applied in Experiment 3c to test the effect of evaluative congruency in this variant of the change detection task without the strong focus on valence. Although the three experiments used slightly different methods, we again chose the exposition as a single study, especially since results were largely comparable.

4.1.1 Common design and procedure

Participants. In Experiments 3a to 3c, 205 students (148 females, age range 18-35 years, *Md* = 23) participated and were paid 4 to 10 € for participation. The data of 4 further participants (from Experiment 3a) were excluded from the analysis because they did not perform significantly above chance in the identity change detection task. All of the participants had normal or corrected-to-normal vision. Table 1 provides a detailed description of the demographic data for the three samples. A power analysis using GPower (Faul et al., 2007), assuming a small to medium effect size of $d_Z = .30$ (above the one observed in Experiment 2), revealed that testing $N = 70$ participants would result in a power of $1\text{-}\beta = .80$.

Table 6 *Demographic data of the samples in Experiment 3a-c*

	Experiment 3a	Experiment 3b	Experiment 3c
N	65	70	70
Gender (f/m)	46/19	56/14	46/24
Median age (Range)	24 (18-32)	23 (18-34)	23 (18-35)

Design and procedure. In each experiment, evaluative congruency was manipulated within participants (congruent vs. incongruent). A change detection task (with change and no change trials) was used. The study array contained two faces, one face on the left and one face on the right. In a given trial, one face was the designated target. (As in Experiment 2a-d, the target status was not evident for the participants.) Whether this face was presented on the right or the left side of the study array was random. In congruent trials, the distractor face expressed the same emotion as the target. Accordingly, in incongruent trials, the distractor face expressed the other emotion. In each of the three experiments, sets with angry and happy faces were used. In half of the trials the target face was happy, and in the other half, it showed an angry facial expression. In change trials, an identity expressing the same emotion as the target was randomly selected to replace the target in the test display. At test and after short retention, the two faces were presented on different locations compared to the study array: They were presented one above the other. It was randomly assigned which of the two faces in the test display was presented at which location. In Experiment 3b and 3c, participants indicated changes by pressing the C-key on a German standard QWERTZ keyboard and when there was no change they had to press the M-key. In Experiment 3a, the S-key and L-key were used and the participants' numbers were utilized to determine which key indicates changes and which key indicates that no change occurred.

Participants were tested individually on personal computers and seated at a distance of approximately 70 cm in front of the screen. Every single face in the learning display covered 2.82° × 3.95° visual angle and they were presented with a horizontal distance of approximately 0.44° visual angle. In the test display, the stimuli had the same size as in the learning display (2.82° × 3.95°) but they were presented one above the other at a distance of 0.52° visual angle. For each of the three experiments, 128 relevant trials entered the main analysis. There were minor procedural differences between the experiments that could be assumed to cause different patterns of results (which was however not the case). One difference is that in Experiment 3a, prior to the presentation of the learning display containing two faces, another learning display showing two abstract shapes was presented. Also after the test display that showed two faces (which were arranged differently than in the learning display), a second test display was presented. This display again showed two abstract shapes and participants had to respond whether these two shapes were the same as the two stimuli at the beginning of the trial or not. This procedure was only used in Experiment 3a. Another difference concerns the use of a change detection task for valence in additional trials (one-third of the total amount of trials). That is, in Experiment 3a and 3b (but not in Experiment 3c), there were trials in which instead of a test display showing two faces a frame was presented at the position of one of the previously studied faces. In this frame, “pos” as an abbreviation for positive or “neg” for negative was written. In these trials participants indicated via key press whether the abbreviation matched the valence of the face that was previously presented on this position or not. The practice phases in each experiment also differed depending on the task characteristics. Besides these differences, presentation time across the experiments was adjusted. More precisely, while in Experiment 3a the study array was presented for 500 ms, the study array in Experiment 3b and 3c was presented for 1000 ms. In Experiment 3a, the blank between study and test was presented for 2000 ms. In Experiment 3b and 3c, this blank was presented for 900 ms. Please note that the procedure and sequence of the core trial remained the same (besides this changes in presentation times) in all three experiments. Only the results of the change detection task for face identities, which was in all three experiments nearly identical, was used for hypothesis testing (the detailed procedures of all three experiments are described in Appendix B). The other kinds of change detection tasks (the change detection task for the valence as well as the change detection task for the shapes, which surrounded the change detection task for face identities) did not enter the analyses.

Materials. In all three experiments, 18 images of adult faces, from the Karolinska Directed Emotional Faces databank (KDEF), were used in grayscale cut-out versions (Lundqvist et al., 1998). Only male identities, with happy and angry expressions, were selected. In Experiments 3b and 3c, the 18 identities were split into two sets. For half of the participants, the identities of the subset A were only presented expressing happiness, whereas the 9 identities of subset B were only presented expressing anger. The assignment of subset and emotion was reversed for the other half of participants. This assignment was also intended in Experiment 3a; however, due to a programming error, it was not fulfilled. One half of the participants saw happy faces from Set A and angry faces from Set B. Due to the programming error, however, the other half of participants saw angry faces, as well as happy faces from Set A. This did, however, not influence results and there is no reason why this should have an influence on the congruency effect that was the main focus of our analyses (for more details see the Methods section of Experiment 3a in Appendix B). In Experiment 3b and 3c, each identity of Set A (for example the identities expressing happiness) had a randomly chosen identity in Set B (the set with faces expressing the other emotion, in this example anger) with which it was never presented together. Thereby, it was ensured that identities in congruent trials and in incongruent trials were drawn from total sets of the same size.

4.1.2 Results

Effect of congruency in *d'*. Table 7 shows the mean performance in the congruent and the incongruent condition as well as the difference between these two conditions – the average congruency effect – quantified in *d'*. As in Experiment 2a-d, the *log-linear* correction was chosen to account for relative hit or false-alarm rates of zero or one (see Hautus, 1995). As is evident from the 95%-confidence intervals for the congruency effects of all four experiments (Table 7), Experiment 3b and Experiment 3c showed a significant *reversed* congruency effect with better performance in incongruent compared to congruent trials. Experiment 3a at least numerically indicates the same unexpected effect.

Table 7 *Mean performance (in d') as a function of congruency in Experiment 3a-c*

Exp	*N*	Congruency effect[a] [95% CI]	Congruent Mean *d'* (*SD*)	Incongruent Mean *d'* (*SD*)
3a	65	-0.11 [-0.25 – 0.03]	1.31 (.54)	1.42 (.65)
3b	70	-0.15 [-0.27– -0.03]	2.13 (.67)	2.29 (.61)
3c	70	-0.18 [-0.30 – -0.07]	2.00 (.61)	2.18 (.64)
Total	205	-0.15 [-0.22 – -0.08]	1.83 (0.70)	1.98 (0.74)

[a] Mean congruency effect = *d'*(congruent) – *d'*(incongruent)

Analyzing all three experiments together in a mixed 2 × 3 ANOVA, with congruency (congruent vs. incongruent) as a within-participants factor and experiment as a between-participants factor, revealed an overall (in)congruency effect, $F(1,202) = 16.94$, $p = .001$, $\eta_p^2 = .077$ ($d_Z = .29$), providing evidence for better performance in the incongruent compared to the congruent condition. This effect was not moderated by the factor experiment, $F(2,203) = 0.35$, $p = .705$, $\eta_p^2 = .003$ ($F(2,202) = 43.55$, $p < .001$, $\eta_p^2 = .301$ for the main effect of experiment). In this analysis, the experiments are all weighted equally irrespective of the number of participants (which should not make a difference due to similar number of participants). Nevertheless, we also tested the incongruency effect in the overall sample, in an analysis without the factor experiment and the effect remained significant, $t(204) = 4.16$, $p < .001$, $d_Z = .29$.

Effect of congruency on reaction times. For reaction times, only correct responses were analyzed. Responses that were 1.5 interquartile ranges above the third quartile of the individual RT distribution (Tukey, 1977) or below 200 ms were excluded, these were 4.35% of correct responses, leaving 13794 trials for analysis (77.39 % after the exclusion of false responses, fast and slow responses and – for Experiment 3a – trials excluded due to a programming error). The mean reaction times for the several conditions and the three experiments are indicated in

Table *8*.

Table 8 *Mean response times (in ms) as a function of congruency and change type in the Experiments 3a-c.*

		change		No change	
Exp	*N*	Congruent Mean in ms (*SD*)	Incongruent Mean in ms (*SD*)	Congruent Mean in ms (*SD*)	Incongruent Mean in ms (*SD*)
3a	65	1130 (197)	1159 (221)	1075 (186)	1141 (191)
3b	70	1188 (293)	1227 (299)	1132 (277)	1158 (294)
3c	70	1093 (198)	1122 (216)	1073 (218)	1073 (209)
Total	205	1137 (237)	1170 (252)	1094 (231)	1123 (238)

To investigate the effect of congruency on reaction times, a 2 × 2 × 3 mixed ANOVA with congruency (congruent vs. incongruent) and change type (change vs. no change) as within-participants factors and experiment as a between-participants factor was performed. This analysis revealed a main effect of congruency with significantly faster responses in congruent compared to incongruent trials $F(1, 202) = 42.70$, $p < .001$, $\eta_p^2 = .175$ ($d_Z = .46$). Additionally, there was a significant main effect of the factor change type, $F(1, 202) = 34.20$, $p < .001$, $\eta_p^2 = .145$. Furthermore, there was a significant interaction of the factors experiment and congruency ($F(2, 202) = 3.90$, $p < .022$, $\eta_p^2 = .037$), and also a significant three-way interaction of the factors congruency, change type, and experiment, $F(2, 202) = 5.33$, $p = .006$, $\eta_p^2 = .050$. The factor experiment failed to reach significance, $F(2, 202) = 2.54$, $p = .082$, $\eta_p^2 = .025$. All other effects were also not significant (all $F < 1.42$ and $p > .246$).

To investigate the interaction in more detail, based on the RTs of the individual experiments separate 2 × 2 repeated measurements ANOVA with the factors congruency (congruent vs. incongruent) and change type (change vs. no change) were performed. In Experiment 3a, this analysis revealed a significant main effect of congruency, $F(1,64) = 23.97$, $p < .001$, $\eta_p^2 = .272$, ($d_Z = .61$) with faster responses in congruent ($M = 1103$ ms, $SD = 178$) compared to incongruent ($M = 1150$ ms, $SD = 188$) trials. The main effect of change type was also significant, $F(1, 64) = 4.28$, $p = .043$, $\eta_p^2 = .063$, revealing that participants initiated responses faster in no change trials ($M = 1109$, $SD = 181$) compared to change trials ($M = 1145$, $SD = 201$). Further, there was a significant interaction of the factors congruency and change type, $F(1, 64) = 4.50$, $p = .038$, $\eta_p^2 = .066$. Decomposing this interaction revealed that, in change trials, response latencies were significantly faster in congruent trials ($M = 1130$, $SD = 197$) than in incongruent trials ($M = 1159$, $SD = 221$), $t(64) = 2.08$, $p = .041$, $d_Z = .26$. In no change trials, this effect of faster responses in congruent trials ($M = 1075$, $SD = 186$) compared to incongruent trials ($M = 1142$, $SD = 186$), was even stronger, $t(64) = 5.21$, $p < .001$, $d_Z = .65$.

The same analysis revealed for Experiment 3b a significant congruency effect, with faster responses in congruent ($M = 1160$, $SD = 280$) compared to incongruent trials ($M = 1192$, $SD = 291$), $F(1, 69) = 12.88$, $p = .001$, $\eta_p^2 = .157$ ($d_Z = .429$). This effect did not significantly interact with the factor change type, $F(1,69) < 1$. There was a main effect of change type, with faster response latencies in no change ($M = 1145$, $SD = 280$) compared to change trials ($M = 1207$, $SD = 293$), $F(1,69) = 30.66$, $p < .001$, $\eta_p^2 = .308$.

In Experiment 3c, the analysis revealed a significant main effect of evaluative congruency with faster response latencies in congruent (*M* = 1083, *SD* = 202) compared to incongruent trials (*M* = 1098, *SD* = 207), $F(69) = 5.80$, $p = .019$, $\eta_p^2 = .078$. ($d_Z = .29$) that was significantly moderated by the factor change type, $F(1,69) = 5.92$, $p = .018$, $\eta_p^2 = .079$ ($F(1,69) = 11.51$, $p = .001$, $\eta_p^2 = .143$, for the factor change type, with faster responses in no change trials [*M* = 1073, *SD* = 212] compared to change trials [*M* = 1108, *SD* = 203]). Further investigation of the interaction revealed that in change trials responses were faster in congruent (*M* = 1093, *SD* = 198) compared to incongruent trials (*M* = 1122, *SD* = 216), $t(69) = 2.91$, $p = .005$, $d_Z = .35$, while for no change trials, there was no significant effect of congruency, $t(69) = 0.03$, $p = .972$, $d_Z = .00$, with *M* = 1073, *SD* = 218 in congruent no change trials and *M* = 1073, *SD* = 209 in incongruent no change trials.

Angry face benefit. The angry face benefit in *d'* was analyzed with a 2 × 4 mixed ANOVA with the within-participants factor emotion (only angry faces vs. only happy faces) and the between participants factor experiment. This analysis revealed a significant main effect of the factor emotion with better performance in trials with only angry faces than in trials with only happy faces, $F(1.202) = 13.99$, $p < .001$, $\eta_p^2 = .065$ ($d_Z = .26$). This effect did however interact with the factor experiment, $F(2, 202) = 3.23$, $p = .042$, $\eta_p^2 = .031$, see Table 9 ($F(2,202) = 34.96$, $p < .001$, η_p^2. = .257 for the main effect of the factor experiment).

Table 9 *Mean performance (in d') as a function of face valence (angry face benefit) in the Experiments 3a-c*

Exp	*N*	Angry face benefit[a] [95% CI]	Angry faces Mean *d'* [b] (*SD*)	Happy faces Mean *d'* [c] (*SD*)
3a	65	0.29 [0.10 – 0.48]	1.46 (.68)	1.16 (.64)
3b	70	0.28 [0.11 – 0.45]	2.27 (.75)	1.99 (.76)
3c	70	0.01 [-0.17 – 0.19]	2.01 (.77)	2.00 (.65)
Total	205	0.19 [0.09 – 0.29]	1.92 (.81)	1.73 (.79)

[a] Mean angry face benefit = *d'*(only angry faces) – *d'*(only happy faces)

[b] performance in angry only trials

[c] performance in happy only trials

To analyze response times, a 2 × 2 × 3 mixed ANOVA with the within-participants factors emotion (only happy faces vs. only angry faces), change type (change vs. no change) and the between-participants factor experiment was conducted. This analysis revealed a tendency for faster responses in congruent trials with only happy faces compared to trials with only angry faces, $F(1, 202) = 3.24$, $p = .074$, $\eta_p^2 = .016$ ($d_Z = .13$). This effect interacted with the factor experiment, $F(2,202) = 3.88$, $p = .022$, $\eta_p^2 = .037$. To further investigate the interaction it was tested for all three experiments separately whether there was a difference in reaction times for trials with only angry and trials with only happy faces. In Experiment 3a, responses were significantly faster in trials with only happy faces ($M = 1090$, $SD = 172$) compared to trials with only angry faces ($M = 1116$, $SD = 195$), $t(64) = 2.31$, $p = .024$, $d_Z = .29$. In Experiment 3b, there was only a tendency for faster responses in trials with only happy faces (*M* = 1151, *SD* = 274) than in trials with only angry faces (*M* = 1169, *SD* = 291), *t*(69) = 1.90, *p* = .061, d_Z = .23, and in Experiment 3c, there was no significant difference in response times in trials with only happy faces (*M* = 1089, *SD* = 207) and trials with only angry faces (*M* = 1077, *SD* = 206), *t*(69) = 1.21, *p* = .232, d_Z = .14.

Additionally, there was a significant main effect of the factor change type, *F*(1, 202) = 27.58, *p* < .001, η_p^2 = .120 (d_Z = .37) with faster responses in no change trials (*M* = 636, *SD* =230) compared to change trials (*M* = 687, *SD* = 240). This effect of change type did not interact with the factor experiment, *F*(2,202) = 2.04, p = .113, η_p^2 = .020, but there was a significant interaction of the factors change type and emotion, *F*(1,202) = 19.21, *p* < .001, η_p^2 = .087 (d_Z = .31). To decompose this interaction the effect of the emotion was investigated for change and no change trials separately. For change trials, there was no significant difference in response times for trails with only angry faces (*M* = 11130, *SD* = 249) and trials with only happy faces (*M* =1144, *SD* = 240), *t*(204) = 1.74, *p* = .084, *d* = .12. For no change trials, responses were significantly faster in trials with only happy faces (*M* = 1076, *SD* = 232) compared to trials with only angry faces (*M* = 1112, *SD* = 246), *t*(204) = 4.19, *p* < .001, *d* = .29.

The three-way interaction failed to reach significance (*F*(2, 202) = 1.35, *p* = .261, η_p^2 = .013). The main effect of the factor experiment also did not reach significance, *F*(2, 202) = 2.22, *p* = .111, η_p^2 = .022.

4.1.3 Interim discussion

The analysis of the evaluative congruency effect conducted above led mostly to diverging effects on performance measured with *d'* and response times. That is in the analysis of *d'*, an unmoderated incongruency effect was observed, while the response time analysis (that is however restricted to correct responses only) revealed faster latencies in congruent compared to incongruent trials. This effect on reaction times was observed in each experiment for change and no change trials except for no change trials in Experiment 3c. Accordingly, a drift-diffusion analysis was conducted that has the potential to disentangle effects of greater emphasis on speed or accuracy from effects on performance to gain further insight into the nature of the effects.

4.1.4 Post-hoc Analysis: Drift-Diffusion Modelling

A common phenomenon in psychological research is the speed-accuracy tradeoff: Decisions can be more accurate but slower when accuracy is emphasized over speed, and vice versa. Therefore, observing diverging results with accuracy and response times as described above, or solely analyzing reaction times or accuracy, can under several circumstances lead to misinterpretations (Voss, Nagler, & Lerche, 2013). In this context, drift-diffusion models (Ratcliff, 1978; Ratcliff & McKoon, 2008; Ratcliff & Smith, 2004) provide an opportunity to gain further insight into the underlying cognitive processes of binary decision tasks. Diffusion models are commonly used in experimental psychology to successfully model data obtained from cognitive tasks in many areas such as lexical decision (e.g. Ratcliff, Gomez, & McKoon, 2004), perceptual discrimination (Voss, Rothermund, & Voss, 2004), priming (Voss, Rothermund, Gast, & Wentura, 2013), as well as a tremendous amount of other tasks (see Wagenmakers, 2009). Most important for the current purpose, drift-diffusion models have also been applied to memory paradigms (e.g. Spaniol, Madden, & Voss, 2006) and also to several working memory tasks (e.g. Pearson, Raškevičius, Bays, Pertzov, & Husain, 2014; Pearson et al., 2014; van Vugt & Jha, 2011; Yu, Chang, & Yang, 2014; Zhang & Rowe, 2014). Also, the classic article by Ratcliff (1978) focused on tasks like the study-test paradigm and the Sternberg task. These tasks are close relatives of the change detection task, which is currently considered to be a state of the art paradigm to obtain a precise measure of working memory capacity (Brady et al., 2011). Further, drift-diffusion models are commonly used in experimental research to extract parameters which are independent of a potential speed-accuracy trade-off as well as a parameter indicating differences in the emphasis of speed or accuracy (e.g. Liu & Watanabe, 2012; Zhang & Rowe, 2014, for theoretical considerations see Voss, Nagler, et al., 2013). Using the response time distribution of correct responses as well as the distribution of errors, these models estimate several parameters that map (1) the information uptake, (2) the amount of information used to make a decision, (3) possible decision biases, and (4) the duration of nondecisional processes (Voss, Nagler, et al., 2013; Voss et al., 2004), which will be described in the following.

(1) The drift rate (*v*) determines the average slope of the information uptake, which describes the speed and the direction of information accumulation (Voss, Nagler, et al., 2013). The drift rate is often suggested to be the parameter indicating the perceptual strength of a stimulus in perceptual learning (Liu & Watanabe, 2012) or quality of evidence being accumulated that is increased after training in perceptual tasks (Zhang & Rowe, 2014), which is independent of a speed-accuracy tradeoff. Applied to working memory tasks, again, a higher drift rate parameter indicates faster evidence accumulation. Formulated differently and applied to the change detection task, if the test and study stimuli are easily distinguishable as different, or if it is easy to recognize that the test items are identical to the studied items, the drift rate is high. For more ambiguous information, the drift rate is rather low (van Vugt & Jha, 2011, for similar arguments see Ratcliff, 1978). In a related working memory task, Pearson and colleagues (2014) provide evidence that the rate of rise of evidence accumulation (μ), which corresponds here to the drift rate (*v*), varies linearly with memory precision. Accordingly, the drift-rate can be interpreted as a measurement of memory strength, which is independent of a speed-accuracy tradeoff.
(2) The threshold-separation parameter (*a*) is a measure of the amount of evidence that is considered for a decision. Low estimates show a liberal decisional style that is characterized by fast but less accurate responses; high estimates of *a* imply accurate but slow responding (Voss, Nagler, et al., 2013). An increase in *a* directly corresponds to a more cautious speed-accuracy setting (Voss, Nagler, et al., 2013). Therefore, for working memory tasks the threshold-separation parameter, *a,* can also be considered to be the parameter that captures the speed-accuracy tradeoff (Pearson et al., 2014; van Vugt & Jha, 2011).
(3) The starting-point (*z*) (or the relative starting point *zr*) reflects the starting point of the evidence accumulation process. It can express response tendencies;

applied to change detection it would mirror a tendency to preferentially respond with change or no change (Voss, Nagler, et al., 2013).

(4) The duration of nondecisional processes (t_0) can express additional processes, for example at encoding or during the response execution, depending on the implemented task (Voss, Nagler, et al., 2013). Applied to working memory or specifically to our paradigm its interpretation is not that clear. It expresses some kind of headstart of processing that is independent of decisional processes, which could be differences in the processing speed of the test faces prior to the actual process of detecting changes. Besides this, the nondecision time also is influenced by other processes not directly linked to decision making like processes of response execution (Voss, Nagler, et al., 2013). It can however not be excluded that retrieval or maintenance processes could also influence this parameter in change detection tasks. A more precisely maintained representation could potentially also provide a headstart for the comparison processes that are needed to decide whether there was a change or not.

For many tasks, there are specific validation studies (e.g. Voss et al., 2004), that is, specific face-valid experimental manipulations were introduced that led to effects on specific parameters (but not on others). For our specific paradigm, no such studies have yet been conducted. Therefore, we must be cautious in interpreting each parameter. This is especially true for the non-decision time $t_{0.}$

Nevertheless, a highly relevant parameter for our purposes is the drift rate *v*. Analyzing the effect of evaluative congruency on the drift rate can reveal whether congruency or incongruency boosts performance independent of a potentially higher focus on accuracy or speed induced by incongruency or congruency. While it is rather sure that this parameter is related to memory performance like precision (Pearson et al., 2014), it is not clear whether it is the *only* parameter that measures precision or whether, for example, the nondecisional processes (t_0) might also be influenced by precision or a better maintenance. Therefore, the nondecisional component t_0 could also be of interest. Its interpretation, however, remains speculative. As described above, a difference in this parameter between congruent and incongruent trials would indicate a headstart before the evidence accumulation in the direction of either a “the same” or a “change” response. Furthermore, differences in the threshold separation may especially be of interest, because they can indicate simple differences in the focus on speed or accuracy.

The previously reported analyses of *d’* and reaction times might give hints about which effects could be expected for the parameters. The faster responses in congruent compared to incongruent trials as well as the better performance in incongruent compared to incongruent trials measured with *d’* might indicate that participants focus more on accuracy in incongruent trials and more on speed in congruent trials. This assumption of a speed-accuracy tradeoff would become apparent in differences of the threshold separation parameters in congruent compared to incongruent trials (*a*). Higher performance in incongruent compared to congruent trials measured with *d’* could suggest that there is a higher drift rate (*v*) in incongruent compared to congruent trials. The differences in reaction times between congruent and incongruent trials could manifest in a difference of the nondecisional component t_0.

Fitting the Diffusion Model to the Data. For drift-diffusion analysis, the removal of outliers is highly recommended (Voss, Nagler, et al., 2013). Therefore, trials with responses faster than 300 ms and responses slower than 5000 ms were excluded from the analysis (0.37% of the data set).

We chose the Kolmogorov-Smirnov method to fit the data as suggested by Voss and colleagues (2004). This approach has the advantage, that it is compared to the log-likelihood statistic not strongly influenced by outliers. Furthermore, and of special importance for experiments with low error rates or low trial numbers (as in the current experiments), no binning of data is required. To fit the data of the individual response time distributions the software *fast-dm* (Voss & Voss, 2007, 2008) was used. To avoid a further reduction of the number of trials on which the parameter estimations are based, for the following analysis, congruent trials with only happy faces and congruent trials with only angry faces were not distinguished between. In the same manner, for incongruent trials, trials in which an angry face was the target (that changed in change trials) and trials in which a happy face was the target were not differentiated between.

The lower threshold was assigned to no change responses, the upper threshold to change responses.

Drift rate *(v)*, the threshold separation parameter (*a*), the non-decision time (t_0) the inter-trial-variability of non-decisional components (*st0*) as well as the relative starting point (*zr*) were estimated separately for congruent and incongruent trials by estimating independent models for these conditions. Further, the drift parameter (*v*) was allowed to vary for change and no change trials. All remaining parameters were fixed to make the model more parsimonious and to improve the stability of the parameter estimates. For statistical analysis, instead of entering the parameters *a* and *zr*, the parameters were reparameterized to calculate the distance of the two thresholds to the starting point: $a_{low} = zr \times a$ (for a no change response), $a_{high} = (1-zr) \times a$ (for a change response).

Statistical analysis of parameter estimates

Drift Rate (*v*). The mean drift rate for the several conditions and the three experiments are presented in Figure 7 to Figure 9. Drift rates were entered in a 2 × 2 × 3 mixed ANOVA with the factors congruency (congruent vs. incongruent) and change type (change vs. no change) as within-participants factors and the between-participant factor experiment (Experiment 1-3). As it could be expected from the analysis of performance measured with *d'*, this analysis revealed a marginally significant effect of evaluative congruency, $F(1, 202) = 3.68$, $p = .057$, $\eta_p^2 = .018$, with a higher drift rate in incongruent ($M = 1.26$, $SD = 0.59$) compared to congruent trials ($M = 1.19$, $SD = 0.60$). This effect was not moderated by the factor experiment $F(2, 202) = 1.27$, $p = .283$, $\eta_p^2 = .012$. The interaction of congruency and change type failed to reach significance, $F(1, 202) = 3.44$, $p = .065$, $\eta_p^2 = .017$. Also the three way interaction did not reveal a significant effect, $F(2, 202) = 1.35$, $p = .260$, $\eta_p^2 = .013$. For the sake of completeness: There was a significant main effect of change type with a higher drift rate in no change ($M = 1.47$, $SD = 0.74$) compared to change trials ($M = 0.99$, $SD = 0.68$), $F(1, 202) = 54.14$, $p <.001$, $\eta_p^2 = .211$. The interaction of this effect with the factor experiment failed to reach significance, $F(2,202) = 2.79$, $p = .064$, $\eta_p^2 = .027$. The factor experiment did reveal a significant effect, with a mean drift rate of $M = 0.83$ ($SD = 0.41$) in Experiment 1, $M = 1.41$ ($SD = 0.48$) in Experiment 2 and $M = 1.41$ ($SD = 0.50$) in Experiment 3, $F(2, 202) = 34.37$, $p < .001$, $\eta_p^2 = .254$.

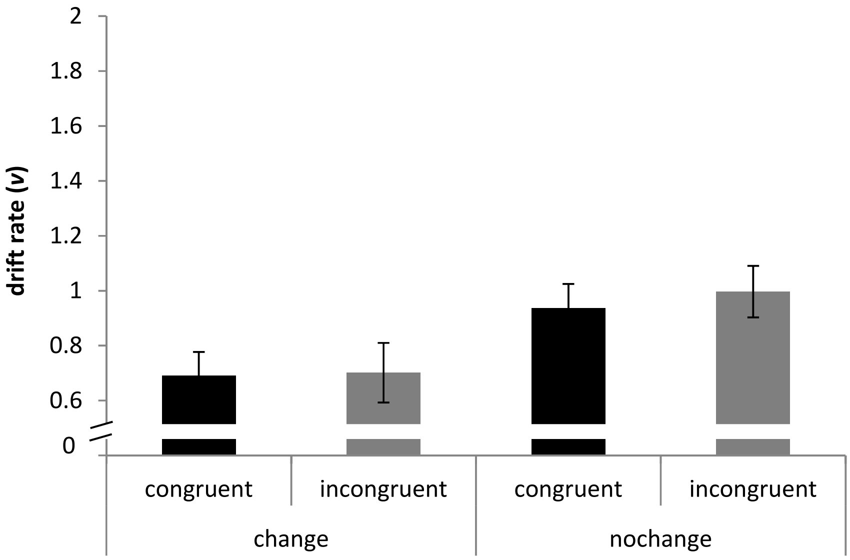


Figure 7. *Mean drift rates (v) as a function of both congruency and change type in Experiment 3a. Error bars represent plus and minus one standard error of the mean.*

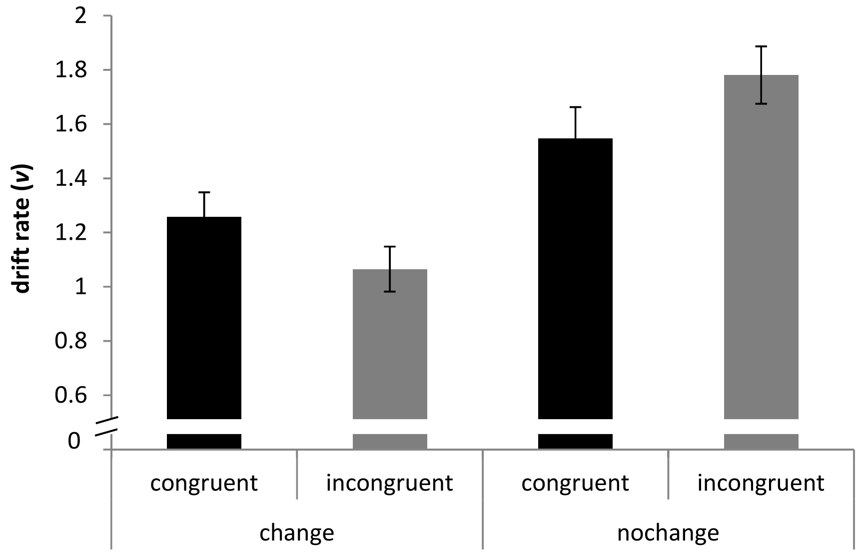


Figure 8. *Mean drift rates (v) as a function of both congruency and change type in Experiment 3b. Error bars represent plus and minus one standard error of the mean.*

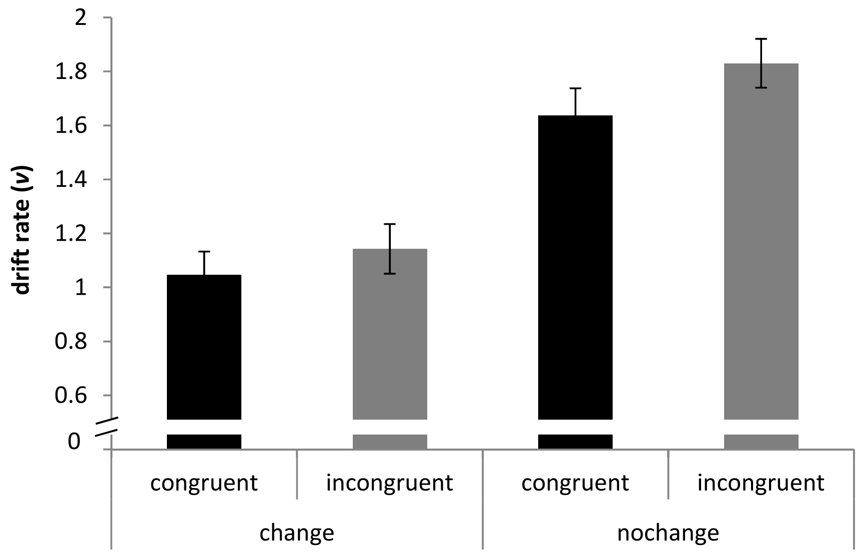


Figure 9. *Mean drift rates (v) as a function of both congruency and change type in Experiment 3c. Error bars represent plus and minus one standard error of the mean.*

Threshold-separation parameter (a_{low}, a_{high}). The mean threshold-separation parameters for the several conditions and the three experiments are presented in Figure 10 to Figure 12. The threshold-separation parameters were entered a 2 × 2 × 3 ANOVA with the within-participants factors congruency (congruent vs. incongruent), threshold (low: no change vs. high: change response) and the between-participants factor experiment (Experiment 1-3). This analysis only revealed a significant main effect of congruency with a higher threshold separation in incongruent (M = 0.77, SD = 0.19) compared to congruent trials (M = 0.74, SD = 0.18), $F(1, 202) = 6.36$, $p = .012$, η_p^2 = .031. All other effects in this analysis were not significant.[20]

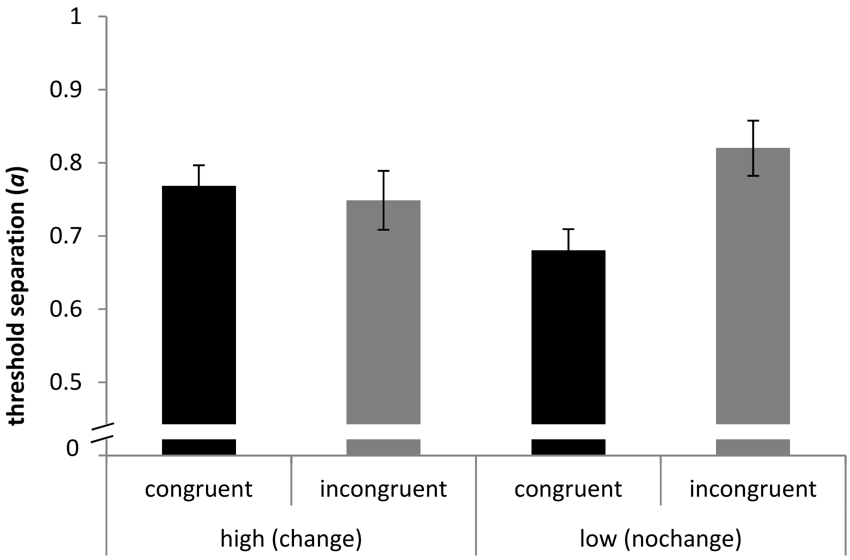


Figure 10. *Mean threshold separation (a) as a function of both congruency and threshold in Experiment 3a. Error bars represent plus and minus one standard error of the mean.*

[20] Two way interaction of threshold and congruency: $F(1, 202) = 2.94$, $p = .088$, $\eta_p^2 = .014$; three-way-interaction: $F(2, 202) = 2.53$, $p = .082$, $\eta_p^2 = .024$, all other Fs < 1.79, all other ps > .17.

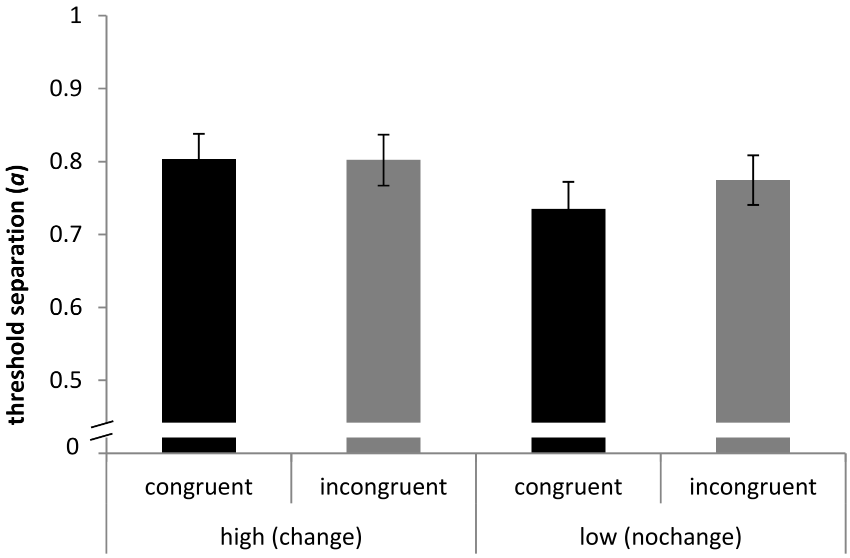


Figure 11. *Mean threshold separation (a) as a function of both congruency and threshold in Experiment 3b. Error bars represent plus and minus one standard error of the mean.*

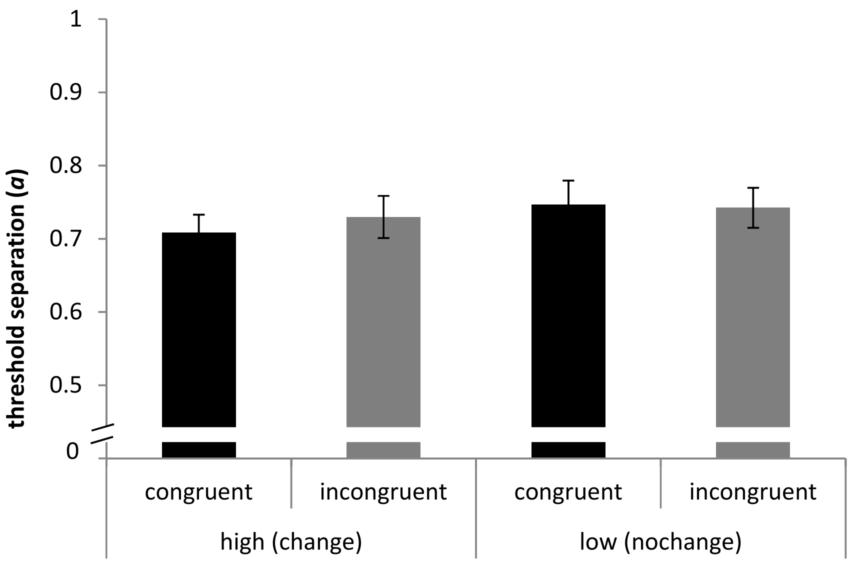


Figure 12. *Mean threshold separation (a) as a function of both congruency and threshold in Experiment 3c. Error bars represent plus and minus one standard error of the mean.*

Nondecisional component (t_0). To analyze the effect of evaluative congruency on nondecisional processes a 2 × 3 mixed ANOVA with the within-participants factor congruency (congruent vs. incongruent) and the between-participants factor experiment (Experiment 1-3) was performed (see also Table 10). A significant main effect of congruency revealed a shorter nondecision time in congruent (*M* = 0.75, *SD* = 0.15) compared to incongruent (*M* = 0.77, *SD* = 0.16) trials, $F(1, 202) = 7.59$, $p = .006$, $\eta_p^2 = .036$. This effect was not moderated by the factor experiment, $F(2, 202) < 1$, $p > .90$. There was also a significant main effect of the factor experiment $F(2, 202) = 4.17$, $p = .017$, $\eta_p^2 = .040$, with a mean nondecisional component of *M* = 0.73 (*SD* = 0.14) in Experiment 1, *M* = 0.80 (*SD* = 0.17) in Experiment 2, and *M* = 0.75 (*SD* = 0.13) in Experiment 3.

Table 10 *Mean and standard deviations of the nondecisional component (t_0) for congruent and incongruent trials in the Experiments 3a-c*

		Experiment 3a	Experiment 3b	Experiment 3c
congruent	mean	0.71	0.78	0.74
	SD	0.14	0.18	0.12
incongruent	mean	0.74	0.81	0.75
	SD	0.15	0.17	0.15

Fit. Table 11 shows the fit-indices (*p*) provided by *fast-dm* (Voss & Voss, 2007). The *p* values are probabilities of the Kolmogorov-Smirnov (KS) statistic. Therefore, a low value indicates a deviance of the empirical reaction time distribution from the predicted RT distribution. For our analysis, the *p* values are a product of two different *p* values, the one for change and the one for no change trials. These *p* values cannot be interpreted as the precise probability of a statistical test. Nevertheless, as can be seen in Table 11 the values are near one. Therefore, the empirical distribution is reproduced rather closely by the predicted distribution. Please note, however, that for small numbers of trials the power of this test might be too small to reliably detect misfit (Voss, Voss, &

Lerche, 2015). On the other hand, in our analysis, the *product* of the *p* values for change and no change trials is calculated by *fast-dm*. This leads to the opposite, making *p* values smaller and thus potentially indicating misfit even when a single KS statistic for a single condition is not significant (Voss et al., 2015).

Table 11 *Mean and standard deviations of the estimated fit for the Experiments 3a-c reported separately for the models for congruent and incongruent trials*

		Experiment 3a	Experiment 3b	Experiment 3c
congruent	mean	.85	.91	.90
	SD	.16	.11	.12
incongruent	mean	.89	.91	.91
	SD	.12	.11	.10

4.1.5 Discussion

For the individual experiments, reaction times and accuracy measured with *d'* indicated effects of congruency pointing in different directions. Therefore, based on these two measurements there is neither clear evidence for or against any specific theory that can make predictions on the effects of evaluative congruency in working memory. However, there is a clear statement that can be made about the results: They are rather incompatible with the finding of Experiment 2a-d, in which there was only an enhancing effect of evaluative congruency observed with *d'*. Thus, assuming only spreading activation processes, or that the formation of compound-cues operates precisely in the same way as we assumed them to in Experiment 2a-d, seems inappropriate. Due to this, we focus here not on a discussion of spreading activation theories or compound-cue theories (besides the retrieval account that might lead to different assumptions for Experiment 2a-d and Experiment 3a-c). Instead, other processes are discussed, that seem to have differentially influenced the results in these two series of experiments, which directly leads to the differences between the used paradigms.

Due to the fact that there was no specific focus on valence in Experiment 3c, and results in this experiment were highly comparable with the results of Experiment 3a-b in which this focus was implemented, this factor seems not to have a strong influence in the current design. Thus, the potentially influential differences between Experiment 3a-c and Experiment 2a-d are, firstly, the set size and, secondly, the change of locations from study to test that was implemented in all three Experiments 3a-c, but not in Experiment 2a-d.

Set size. The first difference to Experiment 2a-d is the set size. When there is an effect of evaluative congruency with set size four but a reversed effect with set size two (at least for *d'*) a more holistic processing might be of importance to obtain beneficial effects in the case of congruency. In a display with four faces arguably, the whole arrangement could be processed at one glance instead of attending each of the four complex stimuli rather separately. In a display with only two faces (especially when their locations change from study to test) the stimuli might be processed more in a face-by-face manner. Although this assumption is rather speculative, the absence of an enhancing effect of congruency in Experiment 2b.2, in the block with upright faces preceded by a block with inverted faces, can also be interpreted in this way. The inverted faces might have reduced holistic processing generally, as the local processing of Navon stimuli can reduce global processing in a different subsequent task (Z. Gao et al., 2011; Liberman & Förster, 2009). While this influence of a reduced global processing mode could be caused by the set size as well as the changing locations from study to test (or even by a combination of both), there might be other processes that are induced only by differences in the set size. Whether such influences exist or not is also investigated and discussed in Experiment 4 in which set size two is used but *without* changing locations from study to test.

Locational changes. As already mentioned another central difference to Experiment 2a-d is the presence of locational changes from study to test in Experiment 3a-c. Participants have to match faces from the test-display to learned faces from the study array because the locations of the stimuli change from study to test. Using the provided information optimally, this can be achieved more easily in incongruent compared to congruent trials: In incongruent trials, the valence of the test faces can be used to know with which of the studied faces a test face has to be compared. A happy test face must only be compared with the studied happy face. The angry test face can selectively be compared to the learned angry face. This is the case because only one identity, but none of the emotional expressions, can change from study to test. However, in congruent trials, there are either two angry faces in both the study and the test display or, alternatively, there are two happy faces at learning and at test. Accordingly, the valence cannot be used to reduce the number of comparisons that have to be performed. Thus, more comparisons are arguably needed to gain insight into whether the stimuli at test are either stimuli previously seen in the study array or not. In other words, the strategy for incongruent trials cannot be used in congruent trials. Therefore, the evidence for a change or a no change decision could be accumulated faster in incongruent trials compared to congruent trials when participants use this strategy. A marginally significant effect in the drift parameter *v* estimated via diffusion model analysis is in line with this assumption. When this process contributes to performance in this design it would counteract beneficial effects of evaluative congruency, for example, due to a mutual facilitation.

There is also a second process that could be introduced as a consequence of the change of locations from study to test. It is possible that the *original* arrangement of the study array was better encoded in congruent trials compared to incongruent trials. This assumption can be derived from the theory by Whittlesea and Jacoby (1990). They suggest that (under specific circumstances) related stimuli (but not unrelated stimuli) form a higher order unit. The change of the locations from study to test potentially destroys this unit that was formed out of the original arrangement, especially when faces expressed the same emotion. This would, therefore, lead to additional costs in congruent compared to incongruent trials. In incongruent trials, in which stimuli are not unitized, they are stored more separately – when the assumptions by Whittlesea and Jacoby (1990) can be applied to the current paradigm. Therefore, destroying the original arrangement has little effect on performance in incongruent trials but it has a strong negative effect on performance in the congruent condition. Whether a task-irrelevant change that destroys the originally stored arrangement and that is similar to the change of the location has an influence on the congruency effect in working memory or not was tested in Experiment 4. Therefore, a more informed discussion of the effect of locational changes is provided there. At this point, it can be concluded that an additional effect might have been introduced by the locational change from study to test, that is, the use of valence to assign test faces to studied identities or the destruction of a unitized arrangement that was formed in congruent trials. In the following section, it is described whether the possible processes suggested here are compatible with the results of the drift-diffusion analysis of Experiment 3a-c.

Discussion of drift-diffusion data. Of more interest than the analysis of accuracy or reaction time data might be the tendency for an effect of evaluative congruency on the drift rate. The drift rate parameter in incongruent trials was higher than in congruent trials, when this marginal effect (p = .057) is taken for granted. At least, it can be stated: If there is any effect of congruency on the drift parameter, evidence accumulation occurs faster in incongruent trials. Therefore, this paradigm seems *not* to measure a mutual facilitation. This finding could, in principle, be better explained by interference in the case of congruency or an enhanced confusability of similar stimuli, in the current task. However, this is an interpretation in contrast to the evidence for better performance in congruent compared to incongruent trials observed in Experiment 2a-d. Therefore, general interference or higher confusability of similar items seems rather implausible, because it remains unclear why it should operate in Experiment 3a-c but not in Experiment 2a-d. Instead of inhibitory processes, other interpretations of this pattern suggest itself: Either reduced set size or the locational changes might have introduced an additional influence, which was not measured in Experiment 2a-d in which congruency led to beneficial effects.

As described above, in incongruent trials, participants might have used the valence of the faces to decide which face of the test display must be compared to which face of the study array. This process cannot be used in congruent trials, where faces cannot be distinguished based on their valence. Therefore, in this way, performance in the incongruent condition might be boosted. This could explain the tendency for a higher drift rate in incongruent trials because the drift rate indicates how fast evidence for either a change or a no change response is accumulated. Thus, when the process described above is used, it is highly plausible that this can be observed in the drift-rate parameter. However, the tendency in the drift rate could also be explained by the retrieval account by Whittlesea and Jacoby (1990). The original arrangement of the study array could be more unitized in congruent compared to incongruent trials. Evidence would then be accumulated faster in incongruent trials because the two stimuli are stored more separately and they can be faster compared to the two differently arranged faces of the test display. Please note that, only a compound-cue theory like the retrieval account (Whittlesea & Jacoby, 1990), which assumes that the formation of a compound depends on the relatedness on stimuli, has the potential to explain the results of Experiment 3a-c and their contrast to the evidence from Experiment 2a-d. However, other compound-cue theories are not capable of explaining the discrepancy between results (M. S. Humphreys et al., 1993; Ratcliff & McKoon, 1988).

Besides a marginal difference in drift rates, the thresholds to give a change or a no change response are less separated in congruent trials, than in incongruent trials. Accordingly, less evidence is considered for a decision in congruent compared to incongruent trials, which means that participants apply a more liberal decisional style, which is characterized by fast but less accurate responses in congruent compared to incongruent trials. Accordingly, participants use different speed-accuracy setting. Explanations of the effect on this parameter remain highly speculative. In incongruent trials, participants might use the valence of stimuli to know which stimulus of the test display has to be compared with which stimulus of the encoding display. This rather elaborated process in incongruent trials could lead to a stronger focus on accuracy in incongruent trials than congruent trials in which this process is not used. This could also cause the difference in the threshold separation between congruent and incongruent trials. When considering an application of the retrieval account to the current paradigm, it remains, however, unclear why a higher unitization in congruent trails (and a following “destruction” of the formed unit) would lead to a relative focus on speed in congruent trials.

Another finding is the longer duration of nondecisional processes in incongruent compared to congruent trials. This effect could be explained, for example, by faster encoding of the congruent items in the test-display. In congruent trials, the two faces in the learning display might prime the encoding of the following two faces in the test display, which express the same emotion. In incongruent trials, the valence in the learning display is mixed. This might not lead to the same amount of proactive priming as in congruent trials. Alternatively, the stimuli in congruent test displays might prime each other's encoding mutually. Therefore, "classical" proactive evaluative priming could arguably explain the difference in the nondecisional component. For many paradigms, it is assumed that proactive priming is mainly a process enhancing the *encoding,* and the non-decision parameter can reveal basic encoding processes (Voss, Nagler, et al., 2013). Another interpretation would be to assume that better maintenance or better retrieval could also manifest in this parameter. Indeed, it is stated that the configuration of working memory can also manifest in this parameter (Voss, Nagler, et al., 2013). More generally, processes that take place prior to the evidence accumulation for the change or no change response (as indicated by the drift rate v) can cause the different parameter estimates (t_0) for congruent and incongruent trials. Therefore, the parameter t_0 could also partially measure the precision of maintained concepts or, more generally, processes during maintenance. For example, while the items in congruent trials might be automatically maintained, in incongruent trials more time might be needed to retrieve the learned items before the evidence accumulation by comparing test faces and the previously learned representations can occur. What effect unitization, which can be assumed based on the assumptions of the retrieval account, should have on the t_0 parameter is unclear. When considering potential effects of the use of valence to assign stimuli of the test display to learned stimuli, it should arguably be assumed that this process is part of the evidence accumulation and not a headstart for the comparison. Therefore, this process would arguably manifest in the drift parameter v instead of the parameter t_0.

To conclude, the effects on most diffusion model parameters seem to be compatible mostly with the use of valence as a retrieval cue in incongruent trials and partially also with an explanation by the retrieval account. The use of valence in incongruent trials as a kind of retrieval cue is compatible with the marginally significant effect on the drift parameter *v*. The effect of evaluative congruency on the threshold separation could also be partially explained by the assumption that this strategy (that focuses on accuracy) is only used in incongruent trials, whereas in congruent trials participants do not rely on this strategy and respond faster. The use of this strategy might however not directly explain the effect on the nondecisional component (t_0). The retrieval account can also predict the effect of congruency on the drift parameter. It is, however, unclear why the formation of a compound should lead to different threshold separations in congruent and incongruent trials. Also, effects of the nondecisional component can arguably not directly be explained by the retrieval account.

Thereby, the suggested processes are generally compatible with the effects on the model parameters. One explanation that is the use of valence as a retrieval cue in incongruent trials makes it plausible that there can be in general a mutual facilitation in the case of relatedness. The (partial) compatibility of the results with the retrieval account by Whittlesea and Jacoby (1990) provides, however, an alternative explanation to a process of a mutual facilitation that cannot be ruled out based on the observed effects. It should, however, be emphasized, that without a precise validation of the parameters for the current task, any of these post-hoc explanations remains rather speculative.

Discussion of the angry face benefit (Experiment 3a-c). Interestingly there was evidence for the angry face benefit in the analysis with *d'*, but analysis of reaction times revealed a partially diverging pattern (in Experiment 3a and 3b or in no change trials). As for the congruency effect, a drift-diffusion analysis seems to suggest itself to compare parameters for trials with only angry faces to trials with only happy faces. Here, we did *not* conduct this analysis, because only half the number of trials could be used compared to the model estimation for the analysis of the congruency effect (here only 32 trials with only angry faces, which are 16 change trials and 16 no change trials, could be used for parameter estimation). Therefore, even the enhanced number of participants that can be achieved by combining all three studies would most likely not compensate for the very small trial number. This makes such an analysis inappropriate.

Accordingly, results of this specific study concerning the angry face benefit remain rather ambiguous. However, taking the results obtained by Jackson and colleagues (2008, 2009, 2014) into account as well as our findings in Experiment 2a-d, the overall impression of a rather reliable and replicable angry face benefit arises. Furthermore, because of the rather low performance in the current study and because the analysis of reaction times is only based on correct responses, it seems likely that an angry face benefit can be assumed for the current paradigm as the analysis of performance measured with *d'* suggests.

Hints for general implications for working memory research. One of the assumptions that could be derived from the findings may be highly relevant for working memory research if it holds true. This is the assumption that participants used the emotions of the faces in the test-display to match the identities of these faces to the identities of the faces from the study array. In other words, we speculated that in the task in which the memory for the identities of faces was tested, participants might have used an independent and task-irrelevant feature (the emotion) to achieve a better change detection performance. If so, this strategy was made useful by changing locations from study to test. In this case, it is obvious that without a change of locations from study to test there would have been no need to use the task-irrelevant dimension in this way to obtain better performance. Therefore, memory for the face *identities* – the task-relevant dimension – would have been measured more reliably when no change of locations from study to test would have been used. If this assumption holds true, whole-display recognition might provide a more precise measure of working memory capacity then single-probed recognition with a centrally presented probe item. That is because the latter task also incorporates a change of the location of the single item in the test display. Therefore, salient but task-irrelevant features (like the emotion of the faces here) might be used as a *cue* that indicates with which item of the learning display the single test item must be compared. That this can boost memory performance becomes evident by a study by Wheeler and Treisman (2002), who observed better performance in whole display recognition when a single relevant item was cued. Following the arguments above, performance in the single probed recognition task with central probes should be boosted by any distinctiveness of the items more than is the case in tasks in which the stimuli (or a single stimulus) in the test display are presented at the same position as before (e.g., whole-display recognition) and in which also perceptual similarity can have a beneficial effect (Lin & Luck, 2009). When a (possibly task-irrelevant) salient feature can be used as a kind of post-cue in single probed recognition with a centrally presented probe, using this procedure in change detection might add a component that resembles the partial report condition of the Sperling task (Sperling, 1960). Thereby capacity (for task-relevant features) might be overestimated when using such procedures. With change detection as a visual working memory task, another legitimate description might be that single probed recognition with a central probe has high similarity to a visual search task, in which the probe is searched in the stored representation of the study array. Obviously, when this is the case, salience and distinctiveness becomes highly relevant and arguably more relevant than when whole-display recognition is used.

With a focus on the congruency effect, however, a logical next step could be an investigation of the effect of evaluative congruency using set size two but no change of locations from study to test. Further, adding a different kind of irrelevant change could be used to test whether either the use of valence as a retrieval cue or the retrieval account can better account for the data of Experiment 3a-c. This was the aim of Experiment 4.

5 Enhanced unitization or mutual facilitation?

In Experiment 2a-d, we observed some evidence for a beneficial effect of evaluative congruency in a working memory task. This is an effect that can be explained parsimoniously by mutual facilitation of evaluatively congruent concepts, which is a process that can also be assumed in priming (Schmitz & Wentura, 2012; Schmitz et al., 2014). Interestingly, in Experiment 3a-c, the opposite was observed: The analysis with *d'* revealed better performance for evaluatively incongruent stimuli in a similar working memory task. Two central procedural differences between the tasks might have caused this difference in results. First, in contrast to our expectations, the reduced set size might have reduced the effect of evaluative congruency, for example, due to a ceiling effect. However, it is hard to imagine how a reversed effect can be explained by this change. Second, changing locations of the stimuli from study to test in Experiment 3a-c might have introduced an additional process. We favored the explanation that participants might have used the emotion of the face stimuli in the incongruent trials of Experiment 3a-c (in which locations of the stimuli changed from study to test) to match the stimuli of the test display to the stimuli from the learning display. Inducing this unwanted strategy can be avoided by using set size two but presenting faces at the same positions at study and test.

However, there is also another potential explanation of the different effects obtained in Experiment 2a-d and Experiment 3a-c. At several places, we referred to an explanation of relatedness effects by "unitization" (Whittlesea & Jacoby, 1990). When evaluative congruency leads to a higher unitization, the alteration of the arrangement from study to test in Experiment 3 by changing the locations might have destroyed this unit, which was formed at learning in congruent but not in incongruent trials. Thereby, costs due to the irrelevant change of locations from study to test might arise in congruent but not in incongruent trials. These costs might not arise when there are no irrelevant changes from study to test.

Applying the retrieval account to predict performance in change detection, the assumptions of this theory can be tested by implementing an irrelevant change paradigm (Ecker, Maybery, & Zimmer, 2013; T. Gao, Gao, Li, Sun, & Shen, 2011; Z. Gao, Li, Yin, & Shen, 2010; Kondo & Saiki, 2012; Zhou et al., 2011). An enhanced unitization of evaluatively congruent concepts should lead to higher costs when the unitized compound is destroyed by task-irrelevant changes. Therefore, other irrelevant changes, and not only the change of locations, could also lead to costs in congruent trials but not in incongruent trials when the assumptions derived by the retrieval account are right. Possibly, this effect might be stronger when stimuli in the learning display are masked or degraded. That is because unitization might be higher for difficult to encode stimuli which raise the need to utilize related concepts (Whittlesea & Jacoby, 1990).

On these assumptions, the next study was built. The study is inspired by research on binding, unitization, and context effects. Please note that in research on binding, two kinds of binding are often distinguished. Due to our focus on the unitization of two distinct items, our approach is more related to the investigation of so-called *inter-item* or *relational* binding or memory but not to research on *intra-item* binding, *conjunctive* binding/memory or binding of features within single object tokens (see Zimmer, Mecklinger, & Lindenberger, 2006, see Mayes, Montaldi, & Migo, 2007 for long-term memory). To follow the considerations above, in the design of the next study in which evaluative congruency was manipulated, it was also manipulated whether there was a task-irrelevant change (different to the change of locations) or not. Thereby predictions by theories assuming mutual facilitation due to a spread of activation and predictions derived from the retrieval account by Whittlesea and Jacoby (1990) can be contrasted. Spreading activation theories predict better performance for evaluatively congruent concepts, but not higher irrelevant change costs in congruent compared to incongruent trials. In contrast, the retrieval account can predict that there is a beneficial effect of evaluative congruency but also higher irrelevant change costs in congruent compared to incongruent trials.

5.1 Experiment 4: The positive effect of evaluative congruency in working memory: Enhanced unitization or mutual facilitation?

In the following study, the hypotheses suggested by mutual facilitation and an application of the retrieval account to working memory were tested. Therefore, a change detection task using emotional faces as stimuli was applied in which evaluative congruency was manipulated. One of the faces was marked as task-relevant (after encoding) for which participants had to indicate whether it changed from study to test or not. The other face was task-irrelevant. Nevertheless, the task-irrelevant stimulus could also change, or not. Thereby, besides a test of a benefit by evaluative congruency (and the angry face benefit) as in Experiment 2a-d and Experiment 3a-c, the costs due to irrelevant changes could also be calculated and compared for congruent and incongruent trials. Finding an overall boost in performance for evaluatively congruent trials in the absence of higher irrelevant change costs in these trials would provide evidence for mutual facilitation. Finding both a higher performance, as well as enhanced irrelevant change costs in congruent trials, would provide evidence for the applicability of the retrieval account to working memory.[21] Furthermore, given that an overall congruency effect will be observed, another question can be addressed: Bearing the results of Experiment 3a-c in mind, it can be stated whether the potential effect of locational changes from study to test in Experiment 3a-c are due to the destruction of a higher order unit or whether the absence of beneficial effects of evaluative congruency in Experiment 3a-c is based on the use of face valence as a cue to match test and study items to ease change detection in incongruent trials. If the latter is the case, there should be *no* significant difference in irrelevant change costs in the following study, in which an irrelevant change different to the change of locations from study to test is used.

---

[21] In the following we use the terms irrelevant change costs and interference index irrespective of the ambiguity of the measurement – that is, it could also indicate a benefit due to the same context at learning and at test. Despite this ambiguity, we use the terms because they are more established in related working memory research (e.g. Kondo & Saiki, 2012).

There is a specific detail in the procedure that we have implemented to achieve a suitable test for our hypotheses. As described above, the authors of the retrieval account state that in priming, when there is a related *degraded* target, the prime combines with the target to form a higher order unit. Hence, they suggest that the prime presented with a degraded, related target is unitized with the target (Whittlesea & Jacoby, 1990). Similarly, in the current task, it might be argued, that it must be comparably hard to initially process the two faces to ensure that, in the case of congruency, both faces form a higher order unit. At least these assumptions suggest that under this condition the use of congruency and unitization could have a higher effect compared to circumstances when the stimuli can easily be processed. Therefore, we decided to use a comparably short presentation time for the two faces of the study array and to use visual masking.

We expected overall better performance in congruent compared to incongruent trials, that can be explained by both theories incorporating spreading activation processes as well as by the retrieval account. Furthermore, we used an index that can be labeled interference index to test whether the irrelevant stimulus is more included into the analysis in congruent compared to incongruent trials, which would be predicted by the retrieval account by Whittlesea and Jacoby (1990).

5.1.1 Materials and Methods

Participants. 67 university students of Saarland University (49 females) with normal or corrected-to-normal vision were paid 4 € for participating. Their age ranged from 19 and 34 with a median age of 25. Data of 3 further participants were excluded because they did not perform significantly above chance. A power analysis based on a small to medium effect size of $d_Z = .35$, with $N = 67$ and $\alpha = .05$ (two-tailed), revealed that a power of $1-\beta = 88.30$ % would be achieved.

Design. As in the previous studies (Experiment 2a-d and Experiment 3a-c) a change detection task was implemented and evaluative congruency was manipulated. In congruent trials, both faces at learning and both faces at test (see below) expressed the same emotion (either happiness or anger). On incongruent trials, a happy and an angry face were presented. In half of these trials the happy face, and in the other half of them the angry face, was task-relevant. As will be described below in more detail, in each trial, one face was marked as task-relevant by an arrow that functioned as a post-cue. The task-relevant face, the target, was the same identity at learning and test in 50 % of the trials. Whether the task-irrelevant face was replaced by another identity expressing the same emotion as before or not was also varied. In 50% of the trials, the irrelevant face changed. These factors, including the emotion of the relevant face (happy vs. angry), the emotion of the irrelevant face (happy vs. angry), the change type of the relevant face (change vs. no change) and the change type of the irrelevant face (irrelevant change vs. irrelevant no change) were varied orthogonally within participants. Whether the face on the right side or the face on the left side was task-relevant was random.

Change-Detection Task.

Figure *13* illustrates the procedure of a single trial. Each trial started with a sequence of a fixation cross ("+") an "X" and again a fixation cross ("+"). This sequence indicated the beginning of a trial, followed by a shortly presented blank. Afterward, a study array was introduced showing two faces, which participants had to remember. After the presentation of two masks and a short retention, an arrow was presented, pointing to the right or to the left. This arrow indicated which face was task-relevant: When the arrow pointed to the left (right), the face that was previously shown on the left (right) position was task-relevant. In the test array, in which again two faces were presented, participants only had to specify whether another face repalced the *task-relevant* face or not. As in the previous studies, the emotions always stayed the same from study to test. Neither the emotions expressed by the faces nor changes in the identity of the task-irrelevant face provided any information that was necessary to solve the task or that could have been directly used to perform better. After a false response, a feedback presenting *Error* (“Fehler”) was presented for 1000 ms. After a correct response, a blank was presented for the same duration (not depicted in

Figure *13*).

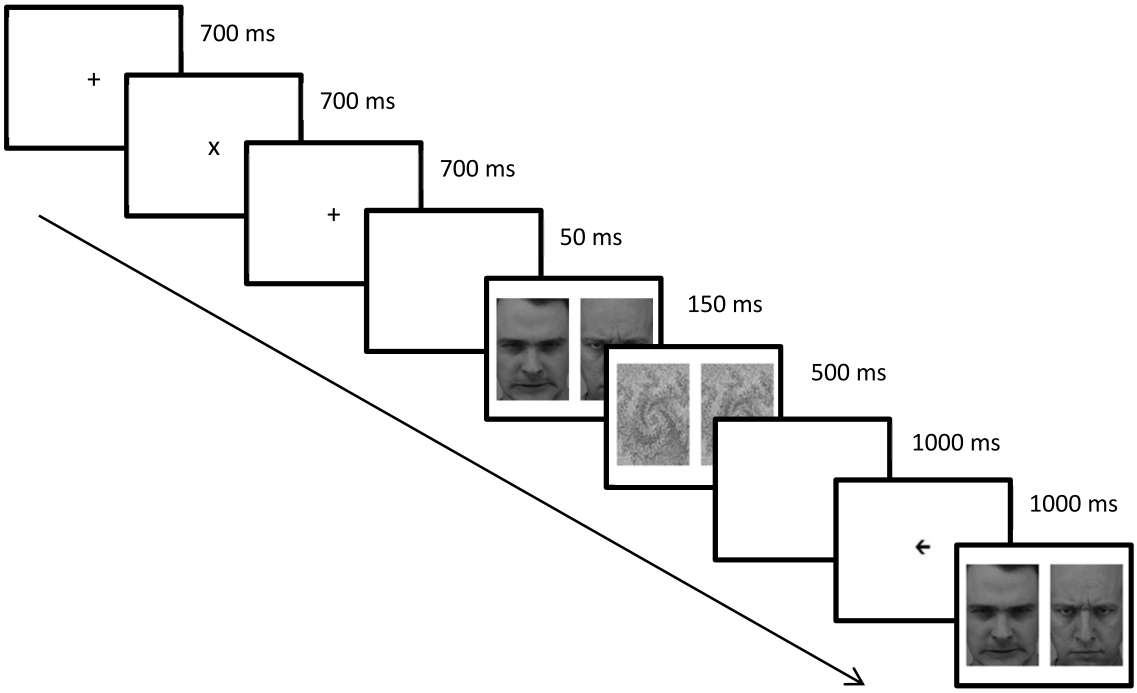


Figure 13. *Trial sequence of the change detection task in Experiment 4. (Depicted is a no change trial).*

The stimuli were 18, grayscale cut-out versions of male faces from the Karolinska Directed Emotional Faces databank (KDEF) (Lundqvist et al., 1998). For each of the 18 selected identities, the expressions happiness and anger were utilized. However, the 18 selected identities were split into two subsets. For nine faces (one subset) only the version expressing anger was shown and for the other nine faces (the other subset) the faces expressing happiness were utilized. This assignment was used for half of the participants. For the other half of partipants, it was reversed. It was further ensured that for each positively valenced face identity one identity was selected that expressed anger with which the positive face was never presented together. Thereby we ensured that there was no different number of possible combinations of face identities in congruent and incongruent trials.

In each trial set size was two. Stimuli were presented at a viewing distance of 70 cm. The two faces shown in the study and the test array covered each 2.49° × 3.48° visual angle. The whole array of both face identities covered 6.38° × 3.48°.

5.1.2 Results

Because irrelevant change costs were calculated, analysis were performed using the percentage of correct responses. Based on the percentage of correct responses for each condition, two planned comparisons were performed. First, the performance in congruent trials was compared with the performance in incongruent trials to test for an overall congruency effect. Second, an interference index was calculated and compared for congruent and incongruent trials. These interference indices were obtained by subtracting the false alarm rate for both-no-change trials (no change of both the relevant and the task-irrelevant face) from the one when (only) the irrelevant stimulus changed (Kondo & Saiki, 2012). In the literature investigating the effect of irrelevant changes, other comparisons are also sometimes reported; however, these contrasts cannot be interpreted unambiguously (Z. Gao et al., 2010).[22] The mean accuracies for the conditions of interest are presented in Table 12.

The first effect of interest was the comparison of performance in congruent trials with the performance in incongruent trials. This test revealed that accuracy was significantly higher in congruent trials (M = 78.75 %, SD = 7.74) than in incongruent trials (M = 76.14 %, SD = 8.47), $t(66) = 3.31$, $p = .002$, $d_Z = .40$. In other words, there was a significant main effect of evaluative congruency. This effect was also confirmed in an analysis using *d'* (*log-linear* corrected, see Hautus, 1995) that was performed to achieve better comparability of results with the other similar experiments reported in this thesis. In this analysis, again a significantly enhanced performance in congruent trials

[22] Comparing the performance in trials with only a relevant change and both-change trials (with a relevant and an irrelevant change) does not provide an easily interpretable comparison for our purpose. One might argue that in both kinds of trials (only relevant change and both-change trials) the potential unit of both stimuli will be destroyed, not yielding a meaningful comparison for our purpose. Therefore, only the interference index described above was used.

(*M* = 1.58, *SD* = 0.48) compared to incongruent trials (*M* = 1.43, *SD* = 0.51) was observed, $t(66) = 3.07$, $p = .003$, $d_Z = .37$.

Table 12 *Mean and standard deviations (in parentheses) of the percentage of correct responses of the conditions of interest in Experiment 4*

		relevant no change		relevant change	
		irrelevant no change	irrelevant change	irrelevant no change	irrelevant change
congruent	happy only	82.28 (15.86)	75.93 (17.18)	75.00 (20.06)	77.80 (19.68)
	anger only	83.21 (13.51)	74.25 (16.84)	78.92 (17.02)	82.65 (18.33)
incongruent	happy target	77.24 (19.57)	68.47 (18.38)	68.10 (21.35)	76.68 (17.12)
	angry target	81.34 (16.19)	76.87 (19.00)	77.61 (18.91)	82.84 (16.40)

To investigate the effect of irrelevant changes, we compared the interference index for congruent and incongruent trials. These indices were obtained by subtracting the false alarm rate of both no change-trials (neither the relevant nor the irrelevant stimulus changed) from the false alarm rate when (only) the irrelevant stimulus changed (Kondo & Saiki, 2012). However, comparing the irrelevant change index for congruent trials (*M* = 7.65 %, *SD* = 14.55) with the irrelevant change index in incongruent trials (*M* = 6.62 %, *SD* = 16.75), no significant difference was observed, $t(66) = .41$, $p = .680$, $d_Z = .05$. Although, there were no differences in the interference indices, the interference indices were significantly above zero in both the congruent condition, $t(66) = 4.30$, $p < .001$, $d = .53$, as well as the incongruent condition, $t(66) = 3.24$, $p < .002$, $d = .40$.

Angry Face Benefit. The analysis revealed no significant difference in percentage correct between displays only showing angry faces (*M* = 79.76%, *SD* = 8.90) and displays only showing happy faces (*M* = 77.75 %, *SD* = 9.89), $t(66) = 1.54$, $p = .130$, $d_Z = .19$. Also in the analysis with *d'*, there was no significant difference in performance between trials with only angry (*M* = 1.64, *SD* = 0.53) and only happy faces (*M* = 1.53, *SD* = 0.59), $t(66) = 1.32$, $p = .191$, $d_Z = .16$.

5.1.3 Discussion

Our results show a clear significant performance enhancing effect of evaluative congruency using a standard working memory task. In the current change detection task, participants performed better if two to-be-remembered faces expressed the same emotion (both faces showing anger or both showing happiness) compared to conditions in which two faces expressed different emotions (one angry and one happy facial expression were shown). This finding replicates the effect obtained in a previous series of experiments (Experiment 2a-d) which was however not observed in another series in which a closely similar task was implemented, but in which the locations of the items were changed from study to test (Experiment 3a-c). As in the previous studies, participants only had to remember face identities, with the emotion of the faces not being task-relevant.

A special characteristic of the current task was that during maintenance one of the two faces was tagged as task-relevant. Following this, participants had to indicate whether the face on the marked position was replaced by another one or not. Whether the other non-marked face changed was irrelevant to the task. This procedure allowed us to calculate an index that indicates the inclusion of the irrelevant face into the analysis. This so-called interference index showed *no* greater inclusion of the irrelevant face in congruent compared to evaluatively incongruent trials. Therefore, with this task, no evidence was obtained that would indicate that in congruent trials both stimuli are more unitized than in incongruent trials. Such a formation of a higher order unit in the case of relatedness can, however, be predicted by applying the retrieval account (Whittlesea & Jacoby, 1990) to the current working memory task. Further, this assumption of a higher unitization in congruent compared to incongruent trials could explain the absence of positive evaluative congruency effects when there are locational changes from study to test as in Experiment 3a-c. Nevertheless, the current data do not provide evidence for this assumption. Instead, the results of the current study are fully compatible with an explanation of the congruency effect by mutual facilitation.

Despite the fact that no direct evidence for the retrieval account was observed in the current experiment, the null effect for the interference index cannot be regarded as sacrosanct evidence against a higher unitization in congruent compared to incongruent trials. Even if the null effect would have been observed with a bigger sample, there might have been several reasons for the lack of the effect.

One reason might be that the need to process the irrelevant stimulus after the onset of the post-cue did not arise in the current paradigm. Although stimuli were masked and performance was relatively low, evaluative congruency might not be regarded as that helpful by the system as stronger semantic relations. Whittlesea and Jacoby (1990) used strongly related word pairs in their study. If the use of evaluatively congruent stimuli is the reason why no effect on the interference index was observed, it can, however, be stated that the retrieval account and the notion of a higher unitization cannot be applied to evaluative congruency in general. Thereby, the retrieval account could only be applied to Experiment 1a-b of the current thesis and to parts of the results by Davelaar and colleagues (2006), but not to the results of Experiment 2a-d, 3a-c and the evidence for mutual facilitation and parallel activation by Schmitz and Wentura (2012, see also Schmitz et al., 2014). Therefore, a mutual facilitation in the case of relatedness seems still to be a more potent theoretical approach, irrespective of this potential critique on the current design. Assuming that evaluative congruency is regarded as helpful and that the assumptions above do not apply, there might be another caveat of the current design. That is, the target defining arrow could have been presented too early or too late. Defining the target too early could cause the irrelevant face to *never* be included into analysis, potentially irrespective of how useful the irrelevant stimulus might be to maintain the relevant stimulus. On the other hand, presenting the post-cue to late might lead to an inclusion of the irrelevant face into the analysis in each trial. However, if a late onset of the post-cue prevented unitization, then there would likely be no unitization in Experiment1a, Experiment 2a-d and Experiment 3a-c either, since there was no post-cue at all in those Experiments. Therefore, unitization could arguably not explain the results of our studies. An early presentation of the post-cue would be even less of a problem because in the original task by Whittlesea and Jacoby (1990) it was always clear from the beginning that one stimulus, the prime, is task-irrelevant. Nevertheless, it might be argued that the time window that was chosen to mark the target in this task did not meet the optimum between these two extremes, in which the irrelevant face *can* but *is not necessarily* included into analysis. Accordingly, the current design might be considered to be a first attempt to unravel influences of relatedness on the irrelevant change effect in change detection. If one of the critiques above does however apply, the retrieval account should also not be capable of explaining the effect of evaluative congruency in this task. That is because the retrieval account assumes that unitization causes the beneficial effect of relatedness. Nevertheless, the null effect should not be over-interpreted.

Essentially, when it is assumed that there is no higher unitization in the case of congruency in this study, it seems also less likely that this process can explain the findings obtained in the similar studies of Experiment 2a-c and Experiment 3a-c. This might even be a small argument against explaining the effect in Experiment 1a and 1b by the retrieval account, especially when it is considered that Experiment 1b (in which a post-cue was also used) has some resemblance to the current experiment.

The, perhaps, more plausible explanation of the current data than differences regarding unitization is to assume that mutual facilitation of evaluatively congruent concepts in working memory led to better performance in congruent compared to incongruent trials. Such a beneficial process in the case of relatedness or congruency can be assumed to take place in working memory according to Schmitz and Wentura (2012) and according to Davelaar and colleagues (2006). This mutual facilitation process can be explained parsimoniously by spreading activation theories (J. R. Anderson, 1976, 1983, 1993; Collins & Loftus, 1975). The theory by Davelaar and colleagues (Davelaar et al., 2006) provides a model in which a spread of activation in working memory causes mutual facilitation. Thus their model is well suited to explain the findings of the current study, in which a classical working memory task showed better performance when evaluatively congruent compared to evaluatively incongruent concepts are maintained. Again, as in Experiment 2a-d, it is also possible that processes merely at encoding (or at retrieval) could explain the effect. However, the finding of a boost in performance due to adding evaluative congruency during the maintenance of valent face-stimuli in a study by Jackson and colleagues (2014) makes this rather implausible.

When mutual facilitation is assumed, the obtained effect size seems plausible. While with set size 4, in which there was also congruency in the incongruent condition, the obtained effect was rather small (d_Z = .17 in the overall analysis of Experiment 2a-d).

Using set size two in the current design a medium effect ($d_Z = .37$, measured with *d'*) was observed. With set size two there is no congruency in the incongruent condition that could also cause mutual facilitation that enhances performance in the incongruent condition.

Alternatively, the evidence in the current study could be still explained by some compound-cue theories different to the retrieval account by Whittlesea and Jacoby (1990). For example, the framework by Ratcliff and McKoon (1988) would predict a higher familiarity signal for stimuli learned in congruent compared incongruent trials, a process that could also explain the overall benefit in congruent compared to incongruent trials. However, this theory has problems accounting for the data obtained in Experiment 1a and 1b, in which the evidence cannot easily be explained by enhanced familiarity. In addition, it can be questioned whether compound-cue theories predict beneficial effects of evaluative congruency and not only effects of stronger semantic associations (Schmitz, 2012).

Overall, when the results from the other experiments are also taken into account, it seems that mutual facilitation of related concepts is one if not the most compelling explanation that can account for beneficial effects of relatedness in both priming and working memory tasks. Nevertheless, some other theories that were not yet applied to semantic similarity or evaluative congruency could explain the current data (see for example Lin & Luck, 2009). That is, for example, a homogenous memory set might produce a kind of "sharpening" (i.e., an increase in precision) of memory representations. This explanation would be in contrast to a mutual facilitation assumption, which is favored here. To conclude, it is up to future research to disentangle between mutual facilitation, compound-cue theories, and sharpening processes.

Angry face benefit. It should be noted that in the current study no significant angry face benefit in working memory was observed. Nevertheless, numerically, performance was better for trials with only angry faces compared to trials with only happy faces. Taking previous research into consideration (in the current thesis and for example the studies by Jackson et al., 2008, 2009, 2014) as well as the non-significant tendency in the current study, it seems like the angry face benefit is still a rather reliable phenomenon.

Previous findings in light of the current data. In the current experiment, performance was increased for evaluatively congruent compared to evaluatively incongruent memory displays. This finding was obtained by implementing a change detection task with set size two. This is in contrast to Experiment 3a-c in which a reversed effect due to evaluative congruency was observed, measured with *d'*. Therefore, the central difference between this study and Experiment 3a-c might have caused the diverging findings, that is, the change of locations from study to test. Taking into account that in this experiment no direct evidence for the retrieval account (Whittlesea & Jacoby, 1990) was obtained, another explanation for the reversed effect in Experiment 3a-c seems promising. That is, the results of the current study suggest that in evaluatively incongruent trials of Experiment 3a-c participants might have used the emotion of the test faces to decide which face of the test display has to be compared with which face of the learning display. More generally, it seems like in change detection with changing locations from study to test, task-irrelevant features can be used similarly to post-cues to ease memory retrieval when the task-irrelevant features generate dissimilarity between items. In this way, change detection tasks in which the item(s) at test are presented at different locations than at learning, might overestimate performance for the task-relevant dimension and lead to wrong estimates of capacity. Thus using single probed recognition with a centrally presented probe could also lead to overestimations of the capacity of visual short-term memory.

6 General Discussion

6.1 Summary of results

The main purpose of this thesis was to investigate the automatic influence of relatedness (semantic relatedness and evaluative congruency) when concepts are concurrently in an active state. These influences were investigated in priming and working memory studies.

Following this idea, after showing semantic priming effects in perceptual identification with an SOA of 0 ms in Experiment 1a, this task was merged with a post-cue task in Experiment 1b. Thereby, a working memory component was introduced into the design. By this means, we were able to provide evidence for the parallel activation of prime and target and a beneficial effect of semantic relatedness when both concepts are concurrently active. Based on the findings and assumptions of Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008) as well as Davelaar and colleagues (2006), the observed beneficial effect of relatedness can be assumed to be caused by a mutual facilitation of related concepts that might arise in working memory. This mutual facilitation of concurrently active concepts can be parsimoniously explained by theories assuming a spread of activation back and forth between related concepts in the semantic network.

A first attempt to investigate the effects of relatedness in working memory by implementing relatedness via evaluative congruency and using change detection provided evidence in line with our assumption (Experiment 2a). Participants achieved higher working memory performance for faces sharing valence in an evaluatively congruent condition compared to an evaluatively incongruent condition. However, first replication attempts have proven difficult. Nevertheless, applying a meta-analytic approach on the highly comparable Experiment 2a-d provided evidence for a small but reliable beneficial effect of evaluative congruency in working memory. With a smaller sample, no effect of evaluative congruency was observed for inverted faces (Experiment 2b.2 and 2c.2). Future research could further use the inversion of faces to test whether the effect of evaluative congruency disappears under this condition. This would indicate that the effect is not based on mere perceptual overlap of facial features. Results can again be explained by a mutual facilitation of concepts that are concurrently active in working memory.

The beneficial effect of evaluative congruency disappeared in Experiment 3a-c, in which set size two and locational changes form study to test were used in an otherwise similar change detection task. More precisely, the analysis of performance measured with *d'* led even to a reversed effect; however, analysis of response times indicated faster responses in congruent compared to incongruent trials. A drift-diffusion analysis revealed that participants focused in these experiments more on speed in congruent trials and more on accurate responses in evaluatively incongruent trials. The analysis of the drift rate parameter revealed a tendency for a higher drift rate in the incongruent compared to the congruent condition. This tendency can be interpreted as better memory performance in the incongruent condition. Based on considerations about possible reasons why no positive effect of evaluative congruency, and even a reversed effect, was observed, a last experiment was designed.

In Experiment 4, again a change detection task with set size two was used. Evaluative congruency was manipulated once again, but this time stimuli did not change positions from study to test. An performance enhancing effect of congruency was expected and observed. Additionally, in this experiment one face was marked as task-relevant by a post-cue. The other face was task-irrelevant. This allowed us to investigate whether there are different effects of irrelevant changes in congruent compared to incongruent trials. This enables comparing the predictions by the assumption of mutual facilitation in the case of relatedness (that can be explained by theories incorporating spreading activation processes) and by another theory predicting beneficial effects of relatedness, namely the retrieval account (Whittlesea & Jacoby, 1990). Evidence was in line with the assumption of mutual facilitation in the case of relatedness, but no direct evidence for the applicability of the retrieval account, which is a compound-cue theory, was obtained. That is, there was better performance in congruent compared to incongruent trials but the effect of an irrelevant change of the non-marked face was not significantly different for congruent and incongruent trials.

The presence of a memory advantage in the case of congruency in Experiment 4 with set size two, makes it plausible that in Experiment 3a-c an additional process was introduced that did not operate in the other studies. This process might have been caused by the change of locations of the stimuli from study to test, which was only present in Experiment 3a-c but not in the other experiments of this thesis. A plausible candidate for such a process is the use of the different emotions of the faces as a cue that helps comparison processes in incongruent but not congruent trials when locations of the stimuli change from study to test. Knowing that faces did not change expressions from study to test, participants might have used the distinct emotions in incongruent trials to know which test face has to be compared with which studied face. In congruent trials in which both faces express the same emotion, this strategy cannot be used. Assuming that this process can explain the results of Experiment 3a-c, the remaining results are all in line with the assumption that related concepts that are concurrently active in working memory do mutually facilitate each other's activation.

6.2 The effects of similarity

6.2.1 Mutual facilitation

The idea that simultaneously active related concepts mutually facilitate each other's activation is theoretically highly plausible. Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008) assume that in priming, prime and target can be active in parallel and – in the case of relatedness – they mutually facilitate each other's activation: The prime helps to encode the target and the target helps to maintain the activation of the prime. They provide evidence for these claims using priming and flanker tasks (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008). They suggest that the mutual facilitation of concurrently active concepts could take place in working memory. This mutual facilitation can be explained by spreading activation theories (J. R. Anderson, 1976, 1983, 1993; Collins & Loftus, 1975; Davelaar et al., 2006). These theories suggest that once a stimulus is activated, activation starts to spread to related concepts. Accordingly, it can easily be assumed that once two related concepts reach the threshold that is needed to start the spread of activation, activation between these two concepts will spread back and forth. The result is an automatic mutual facilitation of related concepts. This assumption was made explicit in the neurocomputational model by Davelaar and colleagues (2006). In this model, a spread of activation leads to a mutual facilitation of related concepts in a limited capacity short-term buffer. Their model can predict various findings in studies testing the memory for word lists.

Also, most experiments in this thesis can be explained by the assumption of mutual facilitation in working memory. Experiment 1b (and partially also Experiment 1a) provides evidence for parallel activation and mutual facilitation of semantically similar concepts in a priming paradigm. By using a post-cue in Experiment 1b, a working memory component is incorporated into the design. Using perceptual identification and controlling for guessing strategies, there are sound hints that the effect is not based on strategic processes but rather on automatic influences of relatedness on performance. The finding, that the effect is higher in Experiment 1b, in which a post-cue was used, than in Experiment 1a, is also in line with the assumption of mutual facilitation in the case of relatedness that takes place in working memory.

Also, the result of the overall analysis of Experiment 2a-d can be explained by mutual facilitation. In a change detection task, evaluative congruency of face stimuli led to a better memory performance. The rather small effect size can be explained by a smaller amount of mutual facilitation in the incongruent condition. While in the congruent condition four faces expressing the same emotion were learned, in the incongruent condition, two happy and two angry faces were presented. Thus there was also (reduced) evaluative overlap in the incongruent condition, which could not have been circumvented in this design with set size four. In line with the assumption of a mutual facilitation due to semantic overlap introduced by evaluative congruency, no effect of congruency was observed with inverted faces for which the processing of the emotion should be reduced. However, the effect for inverted faces was investigated with a reduced sample size. Thus, further research on this issue is needed.

At first glance, Experiment 3a-c, in which a similar design but set size two was used, seems to provide evidence against the notion of a mutual facilitation of evaluatively congruent concepts in working memory. However, the reversed effect in Experiment 3a-c can be explained by the use of valence as a retrieval cue, which was introduced to the design by changing locations of the stimuli from study to test (see section 6.2.7). Therefore, Experiment 3a-c does not necessarily provide evidence against mutual facilitation of evaluatively congruent concepts because the strategy to use valence as a retrieval cue, which counteracts mutual facilitation, arguably circumvented a detection of mutual facilitation processes.

The significant, enhancing effect of congruency in Experiment 4, in which again set size two was used and in which locations of items did not change from study to test, is clearly in line with the assumption of mutual facilitation in the case of relatedness. Furthermore, no evidence for an alternative explanation by a compound-cue theory was provided in Experiment 4.

6.2.2 Compound-cue theories

Compound-cue theories provide an explanation for many priming phenomena that is different from explanations by spreading activation processes. These theories appear to be incompatible with the assumption of a mutual facilitation as suggested by Schmitz and Wentura (2012). In general, compound-cue theories assume that in priming, there is a memory-cue containing the target item and its context that can include the prime (Dosher & Rosedale, 1989; Ratcliff & McKoon, 1988). Therefore, these theories can model the parallel activation of the prime and target concept in priming or of several concepts more generally as a rule. In general, compound-cue theories are theories about the content of retrieval cues and memory access (McNamara, 2005; Ratcliff & McKoon, 1988). Thus, applying or combining these models with working memory models seems plausible at first sight. Further, to explain priming, the compound-cue theory has to be combined with a memory theory. Ratcliff and McKoon (1988) even state that in their theory priming is specifically attributed not to long-term memory but to short-term memory.

Because, in the theory by Ratcliff and McKoon (1988), it is suggested that compounds containing related items lead to a higher familiarity signal, their theory can arguably explain beneficial effects of similarity in the change detection tasks of Experiment 2a-d and Experiment 4. In change detection, higher familiarity for old items compared to new items is essentially sufficient to make a correct change or no-change decision. When this familiarity signal is boosted, especially for learned stimuli in congruent trials but not for learned stimuli in incongruent trials, this could explain the congruency effect. Again, the results of Experiment 3a-c, in which a tendency for a reversed effect was observed for the drift parameter (v), are in contrast to this assumption, but, as we mentioned before, the results of Experiment 3a-c can be explained by the feasible assumption of an additional process. The theory by Ratcliff and McKoon has, however, a problem with other results that are in line with the assumption of mutual facilitation for related items. First of all, the theory arguably has problems accounting for priming effects that are obtained with procedures like the naming task (McNamara, 2005). Therefore, the results of Experiment 1a and 1b of the current study cannot be explained by the compound-cue theory by Ratcliff and McKoon (1988). For the same reasons, their explanation of beneficial effects of similarity, which focusses on differences in the strength of the familiarity signal, seem less well suited to explain the results obtained in the studies by Davelaar and colleagues (2006) or in list-memory paradigms in which free recall is used.

A compound-cue model that is often assumed to overcome the main caveat of the theory by Ratcliff & McKoon (1988), is the theory by M. S. Humphreys and colleagues (1993). They assume that items can generate an associative set. When a representation lies in the intersection of all generated sets (e.g. the intersection of the sets generated by prime and target in priming) it will likely be activated. For semantic priming, it can be assumed that the target will be in the intersection of the sets generated by prime and target when both stimuli are related but not when they are unrelated. Therefore, the model by M. S. Humphreys and colleagues (1993) can easily explain evidence for a parallel activation of prime and target in priming studies using the naming task or the evidence from our priming studies (Experiment 1a and 1b) in which perceptual identification was used. The theory by M. S. Humphreys and colleagues (1993) is also well suited to explain beneficial effects of semantic similarity in memory experiments in which semantically similar words and semantically dissimilar words have to be recalled. Therefore, the theory can also explain the better performance for related words compared to unrelated words in the studies by Davelaar and colleagues (2006). Potentially, it could also explain the zigzag pattern, that the first word of a pair of related words has relatively better performance. The creation of the intersection could reactivate the first word more than the second one because it was forgotten more. Therefore, it would benefit more from the creation of the intersection. In explaining the results of the change-detection tasks of the current thesis, which show better performance for evaluatively congruent compared to evaluatively incongruent concepts (Experiment 2a-d and Experiment 4), the theory, however, has a problem. It is unclear what an associative set could be for face stimuli. Therefore, an application to the working memory experiments of this thesis seems questionable.

Another compound-cue theory is the retrieval account by Whittlesea and Jacoby (1990). Their theory, which was designed to explain semantic priming effects, assumes that performance enhancing effects of similarity occur only under specific circumstances. More precisely, when a target-stimulus is difficult to encode, participants might include a useful, related stimulus into the analysis. Accordingly, only if an irrelevant stimulus is potentially useful (due to its relatedness), and indeed needed (for example due to hard encoding conditions for the target), it is particularly processed. The model predicts that in this case, both stimuli form a higher order unit. In this manner, the retrieval account could be applied to explain much of the current data. In Experiment 1a-b, the perceptual identification task was used. Therefore, there were hard encoding conditions. Thus, according to the retrieval account, there is a need to include useful irrelevant stimuli into the analysis to perform well. Furthermore, the prime is only useful in the related condition but not in the unrelated condition. Accordingly, the retrieval account can predict the positive priming effects that we observed. For Experiment 2a-d, in which a working memory task, namely change detection, was used, there was a high load because participants had to encode and maintain four identities. This might make it hard to perform well in this task. There are trials with evaluatively congruent stimuli in which a higher order unit could be formed to make the task potentially easier and there are other trials with evaluatively incongruent stimuli. Therefore, the overall finding of better performance in congruent compared to incongruent trials could potentially be explained by the retrieval account. However, it remains questionable whether evaluatively congruent faces are considered by the system to be useful for face identification, in the way that words are considered useful for the naming of another related word.[23] Further, another critical point regarding the applicability of the retrieval account to the data of Experiment 2a-d is that all stimuli in the study array are task-relevant. The retrieval account originally assumed that a task-*irrelevant* stimulus would be included more into the analysis in the case of relatedness and when there are difficult encoding conditions. Therefore, it is not explicitly stated whether unitization processes are also triggered when there are only task-relevant stimuli. Although there are unsolved questions, with respect to the results of Experiment 3a-c, the retrieval account is interesting to consider. Let us for a moment assume that evaluative congruency is regarded by the system as useful and that processing is difficult enough to initiate the formation of a higher order unit in congruent trials (but not in incongruent trials). In this case, in Experiment 3a-c, the different arrangement at test might destroy a unitized representation of the two faces that was learned at encoding in congruent trials. Because the arrangement is more unitized in the congruent condition, there should be higher costs due to the rearrangement of stimuli in the test display for congruent compared to incongruent trials. In incongruent trials, in which both stimuli are stored rather separately and independent of each other, the change of the arrangement from study to test should not harm performance very much. Therefore, by changing locations from study to test a process would be induced that leads to a drop in performance in congruent trials but not in incongruent trials. This could explain the tendency for the higher drift

---

[23] Therefore, to explain the findings of Experiment 2a-d with the retrieval account, arguably an additional theory would be needed to explain why and how evaluatively congruent stimuli become useful. Indeed, the usefulness of evaluatively congruent stimuli could be explained by spreading activation theories. It should also be noted that the retrieval account and spreading activation theories might not be as incompatible as it might appear at first glance. The dual-store neurocomputational model by Davelaar and colleagues (2006) that assumes a spreading activation process predicts that in their studies related word-pairs are more likely remembered or forgotten together, a pattern that is indeed compatible with the assumption of higher unitization by Whittlesea and Jacoby (1990).

rate in incongruent compared to congruent trials. To apply the retrieval account to Experiment 4, it is important to note that stimuli were presented for a short duration and masked. These procedural details were intended to pave the way for unitization processes. Furthermore, this experiment was more comparable to a standard priming paradigm because there was a task-relevant as well as a task-irrelevant stimulus. Manipulating change versus no change for the relevant as well as the irrelevant stimulus, we observed that the change of the task-irrelevant stimulus did *not* cause higher costs in congruent compared to incongruent trials. However, this effect (that was not observed) would have provided more direct evidence for the retrieval account: If the unitized retrieval cue that contains both – the relevant as well as the irrelevant stimulus – is destroyed by exchanging the irrelevant stimulus in a congruent trial, there should be more costs compared to the replacement of an irrelevant stimulus that is not unitized with the relevant stimulus (in incongruent trials). By not obtaining this effect, no direct evidence for the retrieval account in working memory is provided. Please note that criticism in such a way that our paradigm did not allow for a higher unitization in congruent compared to incongruent trials would also mean that the retrieval account would not predict a congruency effect. However, such a congruency effect was observed in Experiment 4. Nonetheless, a null effect does not provide ultimate evidence against the applicability of the retrieval account to working memory tasks in which evaluative congruency (or more generally relatedness) is manipulated. Therefore, overall, the retrieval account still offers a promising explanation for some findings of the current thesis, but it needs to be worked out why evaluative congruency should be regarded as useful by the system. In addition, further research is needed to investigate whether similarity can cause unitization when all stimuli are task-relevant.

Further, it is stated that compound-cue theories generally need stronger semantic associations than mere evaluative congruency to explain results (Schmitz, 2012). Additionally, the current compound-cue theories all have some problems in accounting for the evidence for parallel activation and mutual facilitation that was obtained by Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008).

6.2.3 Inhibition and Interference

Overall, the data of the current thesis are rather incompatible with the assumption of a *general* interference, inhibition or a higher confusability in the case of relatedness. Only the negatively singed effect of congruency in Experiment 3a-c (measured with *d'*) could be explained by interference due to evaluative congruency. However, in contrast to this assumption, in Experiment 1a-b, in Experiment 2a-d and in Experiment 4 enhancing effects of similarity were observed. Therefore, taken together, assuming general interference due to similarity in the current studies seems a rather implausible assumption.

However, it could be assumed that semantic similarity and evaluative congruency might trigger inhibitory processes or interference under some conditions. In this case, semantic overlap or evaluative congruency can cause a drop in performance leading to more errors or slower responses in related (or evaluatively congruent) trials compared to unrelated (or evaluatively incongruent) trials. For example, in many paradigms like the picture-word interference paradigm interference effects are measured (Caramazza & Costa, 2000; Collina, Tabossi, & Simone, 2013; Schriefers et al., 1990). This is also the case for studies using post-cue paradigms similar to Experiment 1b (Dean et al., 2001; Hocking et al., 2010; G. W. Humphreys et al., 1995). Interestingly, there are indeed experiments that are conceptually roughly similar to Experiment 1b of the current thesis observing interference effects. First, post-cue tasks in which related and unrelated pictures instead of words were used often indicate interference due to semantic relatedness. For picture naming, there is, however, evidence that object attribute integration (like color-form integration) is the source of this interference (see Dean et al., 2001; Hocking et al., 2010). In Experiment 4b by G. W. Humphreys and colleagues (1995), in which colored words and color cues were used, neither positive priming nor interference effects were observed. Instead, participants showed only numerically faster naming latencies and fewer errors in the related condition compared to the unrelated condition. Still, in their study, object-attribute integration could have influenced the results because the color of the words was used to indicate the target with a post-cue. When it is unlikely that object-attribute integration influences the results because words and rather peripherally presented post-cues that cue a position (instead of a feature) are used positive priming effects are observed (see Dallas & Merikle, 1976a, 1976b as well as Experiment 1b of the current thesis). Therefore, there is no evidence for a general interference between semantically related concepts in paradigms using post-cue tasks. To further validate whether a higher need for an object attribute integration leads also to reversed effects of similarity or evaluative congruency (potentially also in working memory) additional research is needed.[24]

Besides influences of object attribute integration, in the mentioned post-cue studies, differences in task-difficulty could also explain the diverging effects. For Experiment 1b of the current thesis, not only might using perceptual identification instead of naming have introduced a high task difficulty, but also the use of peripheral arrows to mark the target could have enhanced task difficulty. Also for Experiment 4, which is also a kind of post-cue study, task difficulty was rather high. Accordingly, further research could also address how task difficulty might interact with the effect of semantic relatedness in post-cue tasks. For the other working memory experiments of this thesis, task difficulty could potentially also be a moderator for the effect of evaluative congruency. However, it is difficult to compare the studies with this regard because there were several task characteristics that differed.

Please note that we use the terms interference and inhibition here interchangeable. Above, we referred with these terms to potential effects that cause lower performance in conditions with semantic similarity (or evaluative congruency) compared to an unrelated condition (or evaluative incongruency). It should be noted that when better performance in the case of similarity is observed, the current experiments cannot make any statements about whether it is because of a facilitated performance due to relatedness, or because dissimilar (or evaluatively incongruent) concepts inhibit each other. This question can, however, be addressed by future research using neutral conditions. For Experiment 1a and 1b, this could be achieved by using orthographically

---

[24] Perhaps, when the emotion was used by the participants in Experiment 3a-c as a post-cue (in incongruent trials, to know which test item has to be compared with which studied item), this might have led to a higher need for an object attribute integration (the integration of emotion and identity) in Experiment 3a-c compared to Experiment 2a-d and Experiment 4. When this speculative assumption is correct, this could also have contributed to the negatively signed effect of evaluative congruency in Experiment 3a-c.

regular pronounceable nonwords as primes, to calculate approximates of the facilitation due to relatedness and the inhibition for unrelated stimuli (Borowsky & Besner, 1993; McNamara, 2005). In a working memory experiment similar to ours, using evaluatively neural faces could also provide hints as to the degree to which facilitation for congruent stimuli or inhibition for incongruent stimuli causes the pattern of results.

6.2.4 Encoding, maintenance or retrieval processes

It cannot be certainly stated at which stage the advantageous effects of similarity in our tasks arose. For Experiment 2a-d and Experiment 4, the stimuli in the learning display may have mutually primed each other's encoding. Even this process would also be of importance for future working memory research; however, potentially *only* encoding into working memory would be affected. We cannot be sure that mutual facilitation also takes place, once the concepts are maintained in the working memory system. However, there are some arguments against the assumption that processes only at encoding (or only at test) caused the effects that we observed.

First, Experiment 1b, using a post-cue and letting participants maintain two concepts, led to a significantly higher effect size than the fairly comparable Experiment 1a without a post-cue. This is in line with the assumption that the result in Experiment 1b is influenced by mutual facilitation during maintenance, and this effect cannot be easily explained by processes taking place at encoding. Second, in contrast to classical evaluative priming tasks, in Experiment 2a-d, 3a-c, and 4, no direct response to the initially encoded stimuli was required. Thus, it is unclear whether encoding processes can manifest in performance because a completely different task was used compared to priming studies, in which effects are usually attributed to encoding benefits. Third, in a study similar to our experiments (especially Experiment 2a-d and Experiment 4), namely in Experiment 3 by Jackson and colleagues (2014), the authors observed an effect that can best be described as a beneficial effect on memory performance by evaluatively congruent stimuli that were presented during maintenance. More precisely, in their study participants performed a change detection task with valent faces. In the learning display, all faces expressed happiness or all faces expressed anger. During maintenance, a valent word was presented. The match of valence of the learned faces and the word that was presented during maintenance led to better memory performance than when the faces and the word did not match in valence. In other words, in their study, there were no differences between congruent and incongruent trials during the encoding that could have led to the effect on performance. Instead, evaluative congruency introduced during maintenance improved performance. This effect makes it highly plausible that in our design evaluative congruency also improved performance during maintenance. Further reasons to assume an effect during maintenance instead of encoding arise from the research and theories on which the current research is based. The dual-store neurocomputational model by Davelaar and colleagues (2006) suggests that mutual facilitation does occur during maintenance. In their model, the spreading activation process operates in a short-term buffer, which is the activated part of a long-term memory system. This model generates data that match a highly specific empirical data pattern caused by conceptual similarity, namely the so-called zigzag effect. Therefore, mutual facilitation for concurrently active related concepts can explain a very special and otherwise difficult to explain pattern of results.

For Experiment 2a-d and Experiment 4, a potential performance boost due to effects arising from congruency in the *test*-display could also be discussed. The evaluatively congruent faces in the test-display could prime each other's encoding mutually. This could enable faster processing of the test-display. Potentially, this could also lead to a quick comparison of the test display with the maintained representations. However, whether this would explain the effects obtained with *d'* or in accuracy remains unclear. Instead, it should be assumed that such an effect would primarily lead to faster responses. Again, influences at test cannot explain the finding of Experiment 3 by Jackson and colleagues (2014) because they used a single neural test face and evaluative congruency was only induced and present during maintenance. Again, it should be noted that for the study by Davelaar and colleagues (2006) the effects should be attributed to the mutual facilitation of concurrently active concepts, which does not arise at test. Nevertheless, the locus of the effects has still to be addressed by future research.

6.2.5 Perceptual processes

Faces expressing the same emotion are also perceptually similar. Therefore, for Experiment 2-4, there is more perceptual overlap in the congruent compared to the incongruent condition. In Experiment 1a and 1b, in which words were used, the stimuli can obviously be assumed to be equally perceptually similar or dissimilar in the related and the unrelated condition. Thus, the effect in Experiment 1a and 1b can clearly not be explained by perceptual influences. In Experiment 2 of this thesis, in which we also used inverted faces in Experiment 2b.2 and Experiment 2c.2, the effect seems not to be merely explainable by perceptual influences: For inverted faces, which have the same perceptual features but a reduced valence, there was not significantly better performance in congruent compared to incongruent trials. Note, however, that while we analyzed the congruency effect for upright faces in four experiments with a total sample of 207 participants, the effect of congruency for inverted faces was only tested with a total sample of 70 participants. In addition, we did not observe a significant interaction of congruency and inversion when analyzing the experiment in which upright as well as inverted faces were used. Therefore, we only have preliminary evidence that the results in our change detection studies showing beneficial effects of similarity are not based on perceptual similarity. As for mutual facilitation, a potential explanation of the current findings by perceptual similarity does not account the inversion of the effect in Experiment 3a-c compared to the other studies of this thesis.

In all change detection experiments of the current thesis, the degree of perceptual similarity was different for the congruent and the incongruent condition. Nonetheless, there is to our knowledge yet no compelling evidence from studies that show that perceptual influences in our task could provide an alternative explanation for the better performance in congruent compared to incongruent trials. Note that in all studies that observed better performance for perceptually similar stimuli, like similar colors (Lin & Luck, 2009), or for example lines of similar length (Sims et al., 2012), the feature that was tested was the feature that defined the similarity between the stimuli. In the current experiments, however, the valence of the stimuli introduced the similarity, but memory for the identity of the depicted persons was tested. Therefore, in contrast to the studies mentioned above, in our case the features defining similarity and the features that were tested were independent. In any case, whether the current effects are nonetheless based on perceptual similarity or instead, as we assume, on semantic similarity, the current results indicate a performance boost due to similarity in a working memory task in which the similarity defining feature is not task-relevant. This is a finding that was to our knowledge not shown before. To investigate whether mere perceptual similarity of a task-irrelevant feature can improve working memory performance or not, further research is needed.

6.2.6 Theories inspired by the examination of perceptual overlap

While it is implausible that perceptual processes can explain all beneficial effects of similarity in the current thesis, the theories used to explain effects of perceptual overlap seem to also provide promising frameworks to explain other effects of conceptual similarity in a more general fashion. One explanation by Lin and Luck (2009) for beneficial effects of perceptual similarity is that inhibitory interactions might lead to a sharpening of the memory representations in color space. This sharpening takes place when representations of colors lie near to each other. Because of this sharpening, the representations of perceptually similar stimuli (in their case similar colors) are assumed to be activated more precisely. In general, this process could also happen in semantic space. That is, the concept *Bentley* and the similar concept *Porsche* must be more distinguished than the concepts *Bentley* and the concept *suitcase*. This could lead to a more precise activation in the case of relatedness. It is plausible that in the example with related concepts, differences between the two brands might become more accessible and therefore the two concepts could be active more precisely. In the unrelated example one might only know that a car brand was presented, not knowing precisely which one. This would correspond to a less precise representation in semantic space. This explanation could easily explain the results of Experiment 1a and 1b and other priming studies in general. However, to test the applicability of this theory to semantic space, further research is needed. Similarly, a second explanation that Lin and Luck (2009) suggest could be applied to the semantic space. They suggest that for similar representations one representation could serve as an anchor point for another related representation and reduce a drift of the representations with progressing time. Further, a similar representation would provide a better anchor point than a dissimilar representation. Again, this process could also be assumed in semantic space and for priming research. The third explanation suggested by Lin and Luck (2009) is that it is easier to attend to only a small region of color space compared to attending to several regions in color space. Therefore, it should be easier to maintain similar colors than dissimilar colors. As for the processes before, this process might also be assumed to take place in semantic space. Interestingly, assumptions like these were not yet applied to semantic priming or evaluative priming. However, for all of these processes, more research is needed to validate them and to derive distinct predictions that allow testing the theories against each other in working memory research as well as priming research.

Another idea to explain beneficial effects of perceptual similarity on working memory performance was described by Sims and colleagues (2012). They suggest that the rate-distortion theory (Berger, 1971; Shannon & Weaver, 1949) provides a model that can explain why participants store similarly oriented arrows or lines with similar length better than the corresponding dissimilar stimuli. More precisely, they assume that features drawn from a distribution with a lower variance can be stored using fewer bits (pieces of information). Under the precondition that the same assumption could be made for the storage of semantic concepts, this theoretical framework could also be applied to explain effects of semantic similarity in working memory paradigms. Another theory that makes similar assumptions is provided by Brady and Alvarez (2011). As Brady and Alvarez (2011) state, most (visual) working memory theories treat single items as independently stored units with no interactions between items. What they suggest is that higher order information is used to compute statistics of the display. In this manner, they observed that the remembered size of a single item is biased toward the mean size of the items in a display. Furthermore, it is also specifically biased by the mean size of the items in the same color. First, these effects of ensemble encoding show that items are not stored independently. Second, the evidence provides another potential influence of the congruency in Experiment 2a-d, 3a-c and 4 of the current thesis. Participants might "average" on several dimensions over evaluatively congruent items. These dimensions could include visual dimensions (e.g. like the length of the noses) but maybe they also include more semantic dimensions (e.g. femaleness, aggressiveness etc.). Being able to use the statistics of an arrangement to store information might be one process that enables us also to store more information than without using such statistics. Thereby, research on the use of ensemble statistics in visual working memory could also provide valuable and more general ideas to explain beneficial effects of perceptual similarity, evaluative congruency and also semantic similarity on working memory performance. Some of these theories could also be applied to semantic and evaluative priming.

6.2.7 Assignment based on valence

The assumptions explained in the following can clearly only account for a subset of the current experiments. More precisely, only for Experiment 3a-c can this process be assumed to take place. The logic is the following: If stimuli in a change detection task are not presented at the same locations at learning and at test, distinct stimulus features might help the participants to know which stimulus of the test display was presented at which location in the study array. In Experiment 3a-c, in the incongruent condition, the positively valenced face learned at encoding must only be compared with the positively valenced face in the test display, but not with the face expressing anger. This is because the valence of the stimuli did not change from study to test. Correspondingly, the angry face of the test display must also only be compared to the angry face from the learning display. In this way, valence might have been used as a kind of post-cue in incongruent trials. However, in congruent trials with only faces expressing happiness (or only anger), there are no obvious hints as to which positively valenced face in the test-display has to be compared with which positively valenced face from the learning display. Therefore, this process, which is only useful if stimuli change locations from study to test, could trigger a performance benefit in incongruent trials compared to congruent trials. This process would cancel out any effects by mutual facilitation for congruent stimuli. This explanation fits the overall pattern of results that we have obtained in Experiment 3a-c. Importantly, in all reported experiments in which this process should not influence the results (because locations of the stimuli did not change from study to test and in which it was clear which face has to be compared with which face), enhancing effects of evaluative congruency were observed (Experiment 2a-d and Experiment 4).

An important aspect of this argumentation is that even dissimilarity of a task-*irrelevant* feature (in our case valence) could enhance performance in a change detection task in which changes of another, task-relevant stimulus dimension have to be detected. Obviously, if this assumption is true, the same should also count for dissimilarity on the task-*relevant* dimension, when locations change from study to test. Assuming such a process, for which task-irrelevant features can be used, the triggered boost in performance is not necessarily caused by better maintenance or higher precision of the to-be-remembered stimulus or its task-relevant features. Instead, in a change detection task with locational changes from study to test, vague and imprecise but distinguishable representations could also be used to reduce the number of candidates that have to be compared to a specific stimulus of the test display. This could lead to large overestimations of working memory capacity. Whether locations of stimuli change from study to test or not would be a highly potent moderator for effects of similarity on working memory performance.

When such a process is used, it could also occur when a single centrally presented probe is used in change detection. It is a central difference of implementations of the change detection paradigm whether the test display contains a single item, which is often presented centrally, or whether the test display matches the configuration of the learning display – with a change of a single item in change trials (Rouder et al., 2011). When a single centrally presented item represents the test display, participants could use highly salient features of the test item (even if they are task-irrelevant) to compare the test item only to a reduced number of candidates from the learning display. Therefore, this type of single probed recognition has some similarity to the partial report procedure by Sperling (1960). Using a task with single probed recognition as well as centrally presented probes, and assuming that the information by the single probe is used like a cue that limits the number of candidates from the learning display, working memory capacity could be overestimated in such tasks (especially for dissimilar stimuli). The assumption of an artificial performance boost that operates in this way also becomes evident when the research by Wheeler and Treisman (2002) is considered: They observed better performance in whole display recognition when a single item was cued as relevant (for a similar effect of a post-cue see also Makovski, Sussman, & Jiang, 2008; but see Luck & Vogel, 1997). Accordingly, single probed recognition can cause, under some circumstances, severe problems, although one might argue that whole display recognition is also not a "pure" measure of working memory capacity because of influences on performance due to configuration information (Boduroglu & Shah, 2009; Jiang et al., 2000; Logie et al., 2011; Mutluturk & Boduroglu, 2014, but see Woodman et al., 2012). Potentially beneficial effects of several kinds of similarity like evaluative congruency (Experiment 2a-d and Experiment 4 of this thesis), perceptual similarity (Lin & Luck, 2009; Sims et al., 2012) or combined effects of conceptual overlap and perceptual similarity (Jiang, Lee, et al., 2016) should not be replicable using single probed recognition with centrally presented probes. Therefore, the current results might suggest a valuable hint for further research on effects of similarity on working memory performance measured with change detection paradigms.

6.2.8 Global or holistic processing as a moderator?

It can be speculated that in Experiment 3, in which the arrangement changed from study to test, the display was encoded in a face-by-face fashion. In contrast, in Experiment 2a-d with a set size of four and Experiment 4, in which the learning display was masked, the display might have been processed more holistically. Therefore, holistic processing could also determine whether similarity has a beneficial influence on task performance or not. Interestingly, in Experiment 2b, when the effect of congruency was investigated for upright faces that were presented after participants performed the same task for inverted faces, the positively singed congruency effect reversed (see Appendix A). Inverted faces are processed less holistically than upright faces. In addition, it can be assumed that holistic or feature-based processing can be experimentally induced, for example by previously performing local or holistic processing of Navon stimuli (e.g. Gao et al., 2011; Liberman & Förster, 2009). Please note, that it is still an open question whether these inductions can last for longer periods. Following these thoughts, the carry-over effect in Experiment 2b (see Appendix A) could also be explained by reduced holistic processing, when upright faces follow inverted faces. It is however plausible to assume that such effects could be restricted to effects of similarity measured with tasks using visual stimuli. An interesting implication could be that in collectivistic cultures in which holistic perception is more pronounced (Nisbett & Miyamoto, 2005), similarity could increase performance more than for participants from an individualistic culture.

6.3 Angry face benefit

Besides Experiment 1a and 1b, each experiment that was reported in this thesis also provided the opportunity to investigate the angry face benefit in working memory. Here, we analyzed the angry face benefit as the working memory performance in trials with only angry faces compared to the performance for trials with only happy faces (Jackson et al., 2009). Although this phenomenon is highly relevant for emotion research, it is rather unrelated to the previously mentioned theories and research focusing on effects of similarity.

Experiment 2a-d provide clear support for the angry face benefit: Participants showed higher performance for trials in which they had to remember only angry faces compared to trials in which they had to remember happy faces. For Experiment 3a-b, the effect was observed with *d'* but reactions were partially faster in trials with only happy faces. However, for Experiment 3a-b, not only the analysis of the angry face benefit but also the investigation of the effect of similarity revealed diverging results in accuracy and reaction times. Therefore, some task-specific characteristics of this series of studies might have led to the presence of a speed-accuracy trade-off and some unusual processes. Due to the rather restricted number of trials that enter the calculation of the angry face benefit in Experiment 3a-c per participant, no drift-diffusion model analysis was performed (because only half of the trials would be available, as compared to the analysis of the congruency effect). In Experiment 4, we did not observe a significant angry face benefit, but numerically performance was better in trials with only angry faces compared to trials with only happy faces. It should be mentioned that in all of our current studies, the analysis of the angry face benefit is restricted to a reduced number of trials compared to the analysis of the effect of evaluative congruency. The current evidence can be regarded overall as being rather in line with the assumption of better working memory performance for angry faces. Furthermore, it should be kept in mind that a substantial body research by Jackson and colleagues (2008, 2009, 2014, but see Langeslag, Morgan, Jackson, Linden, & Van Strien, 2009) provides evidence for a clear angry face benefit in working memory.

An interesting result of Experiment 2b.2 and 2c.2 is that the angry face benefit was also observed when inverted faces instead of upright faces were utilized. Therefore, in this procedure, the angry face benefit seems to be caused by the perceptual features of the faces. Due to the absence of an effect for inverted faces in a similar study investigating the angry face benefit (Jackson et al., 2009) further research is needed to clarify whether the angry face benefit is caused by the perceptual features of faces or not. It should be noted, that in Experiment 5 by Jackson and colleagues (2009), in which no angry face benefit for inverted faces was observed, half of the participants saw face stimuli that were artificially modified (morphed with neutral faces) to keep the expression intensity constant for angry and happy faces. The other half of participants saw faces from a rather old set by Ekman and Friesen (1976). Potentially, the perceptual features of naturalistic stimuli generally are sufficient to cause an angry face benefit also for inverted faces. Although Jackson and colleagues (2009) used both naturalistic faces as well as modified stimuli, the angry face benefit was not moderated by this factor. Please note, however, that the absence of a significant interaction with the factor set in Experiment 5 by Jackson and colleagues (2009) could be explainable by the rather small sample size.

6.4 Implications and future directions

Based on the effects of similarity in the current thesis, there are three main areas for which implications can arise and in which further research might prove fruitful.

6.4.1 Parallel activation and mutual facilitation in priming

The research by Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008), the theory and evidence by Davelaar and colleagues (2006) and most of the experiments conducted in this thesis are in line with the assumption that simultaneously active related concepts mutually facilitate each other's activation. Especially the research by Wentura and colleagues and Experiment 1b provide evidence relevant for priming research: In priming, prime and target can be active simultaneously and these simultaneously active concepts can mutually facilitate each other's activation in the case of relatedness. Some theories like spreading activation theories can easily account for mutual facilitation and parallel activation in priming.

When discussing parallel activation and mutual facilitation, another highly relevant group of models used to explain priming, the parallel distributed processing models, should be considered with caution. A prominent example is the model by Masson (1991, 1995). In his model, priming effects are explained by a transition of the prime activation pattern into the activation pattern of the target. Concepts are represented by the activation pattern of several nodes and related concepts have similar patterns. Switching activation from one pattern to a pattern of a related concept occurs faster than changing activation to represent an unrelated concept. Therefore, the temporal order of the activation of the concepts is of importance to explain priming in this way. Without additional assumptions, the framework of parallel distributed processing theories can only explain priming effects when the activation of the prime *precedes* the activation of the target. Accordingly, these models are in contrast to the assumptions of parallel activation and mutual facilitation in priming. At least, additional assumptions have to be made to explain the findings of the priming studies by Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008) or Experiment 1b of this thesis by these models. Also, to explain beneficial effects of relatedness or evaluative congruency in working memory like the evidence by Davelaar and colleagues (Davelaar et al., 2006) or Experiment 2a-d and Experiment 4 of the current thesis, recent parallel distributed processing theories need additional assumptions. However, there might be a way to explain priming effects by the model by Masson (1991, 1995) that might not directly be in contrast to the assumption of a parallel activation of concepts in priming. One would have to assume, that a single concept is loaded into a system like the one proposed by Masson (1991, 1995). Within this system, priming can be explained by a faster transition of the activation of one concept into the activation of another concept, when both concepts are related. In this case, the parallel activation could only take place in another *distinct* subsystem, in which representations can be maintained simultaneously; however, similarity would not cause any effect in this system. Furthermore, in these two subsystems, concepts would have to be represented using completely different representational formats. Assuming that several concepts are held active simultaneously while a single concept is in the focus of attention is not implausible, and this assumption is also found in the memory model by Oberauer and colleagues (Oberauer, 2002, 2006, 2009a; Oberauer & Lange, 2009; Oberauer et al., 2013). Still, current parallel distributed processing theories have difficulties because different representational codes would be needed for the system in which only one concept is active (e.g. the focus of attention) in which priming is explained and the other system in which parallel activation of several concepts can be explained. Despite this critique of parallel distributed processing models in this context, it should be noted that these models are neurologically more plausible than spreading activation theories with localist representations of concepts (e.g. Farah & McClelland, 1991). Further, parallel distributed processing models can explain many effects (like the effect of intervening stimuli in priming) that challenge current spreading activation theories (Masson, 1991, 1995).[25] Therefore, creating new versions of these models that

---

[25] At this point, it should be mentioned that there is good reason to assume that spreading activation theories and parallel-distributed processing theories are to some extent compatible. Collins and Loftus (1975) address in their work "features models" (Rips, Shoben, & Smith, 1973; Smith, Shoben, & Rips, 1974). In feature models, concepts consist of a set of values for several semantic dimensions like color, animateness and so on. In recent parallel distributed processing theories like the theory by Masson (1991, 1995) units do not necessarily correspond to namable features, but the basic principle of these theories is the same. Both are theories in which knowledge about a concept is distributed in a network. Therefore, these feature models can to some extent be considered to be instances of parallel distributed processing models. Interestingly, Collins and Loftus state that "Any process that can be represented in a feature model is representable in a network model…" (Collins & Loftus, 1975, p. 410). In a semantic network model like the one by Collins & Loftus (1975), each feature is also a concept, and each concept (like for example the concept cat) can be linked to all its features (like living, elegant, furious, headstrong, …) but it is not linked to unrelated features (like vegetable, flying etc.).

are capable of explaining parallel activation within the same system in which priming effects occur (or at least in a system that does not need another representational format) could lead to new potent theories that could be used to explain an extensive body of evidence in both priming and working memory research.

Please note that there are current working memory models that implement synchronous firing for feature nodes that belong to the same concept (e.g. Raffone & Van Leeuwen, 2001; Raffone & Wolters, 2001; Vogel et al., 2001; Wolters & Raffone, 2008). These are models that can account for parallel activation within a parallel distributed processing model. However, these models currently do not implement mutual facilitation processes. Therefore, Schmitz and Wentura (2012) suggest, that these models could be extended to implement mutual facilitation of related concepts. Such modified theories could be addressed by future research.

Besides theoretical implications, further research can be conducted using several experimental approaches to validate the assumptions of a parallel activation of prime and target and mutual facilitation in the case of relatedness in priming paradigms. First, the research by Schmitz and Wentura (2012) can be further extended. In Experiment 4a/b of their study, no evidence for a mutual facilitation based on semantic overlap different to evaluative congruency was observed. Taking the results from Experiment 1a and 1b of this thesis and the research by Davelaar and colleagues (Davelaar et al., 2005, 2006, see also Haarmann & Usher, 2001; Usher & Cohen, 1999) into account, it is

---

Thereby, a feature model can be implemented within the spreading activation theory by Collins and Loftus (1975) or spreading activation theories in general. When a broader view is applied, this illustrates that also parallel distributed processing models like the one by Masson (1991, 1995) and spreading activation theories (J. R. Anderson, 1976, 1983, 1993; Collins & Loftus, 1975) might be partially transferable into each other. In Masson's model, each node is on or off and the pattern of activation of nodes constitutes the activation of a concept. Similarly, in spreading activation theories, when feature nodes (the concepts of features) are activated, activation would spread to the concept they define (like *cat*). Also, in the other direction, when a concept is activated its feature nodes will become activated by spreading activation processes in spreading activation theories. Taking this similarity of parallel-distributed processing models and spreading activation theories into account, two implications arise. On the one hand, some common criticism of spreading activation theories could be overcome by designing them more similar to parallel distributed processing models. On the other hand, the relatedness of parallel distributed processing theories and classical spreading activation theories provides a hint that recent parallel distributed processing theories could be modified in a way that enables them to account for a parallel activation of several concepts.

highly unlikely that the effect that Schmitz and Wentura (2012) observed is valence-specific. Thus, the design by Schmitz and Wentura could be used with more distinctive semantic categories or with concepts that have a higher semantic similarity within the category (e.g., foods and dogs) than the rather broad categories of persons and animals that were used in the original study. Alternatively, potentially more relevant categories, like the use of living vs. non-living things, could be better candidates for measuring mutual facilitation processes based on another dimension than evaluative congruency. A second line of further research could investigate moderators of the effect of semantic relatedness in the combination of the perceptual identification task and the post-cue task as used in Experiment 1b. As outlined in the Discussion of Experiment 1, manipulations of the delay between the display containing prime and target and the presentation of the post-cue could be used. Following the assumption of spreading activation processes in working memory, the longer the delay, the more activation should add up in the target node leading to a stronger effect of mutual facilitation. Further, the assumption of mutual facilitation in priming could be tested by comparing the effects of symmetrically associated prime-target pairs and prime-target pairs with an association that is primarily from the prime to the target. Assuming a *mutual* facilitation process, for symmetrically associated prime-target pairs, there should be a more rapid lexical-semantic activation-buildup of the target activation. More research on this topic could contribute to the notion that parallel activation and mutual facilitation in the case of relatedness are criteria that theories of semantic priming should be able to account for.

6.4.2 Semantic relatedness in working memory research

Besides the implications for priming research, similar implications for working memory research arise. While priming research can benefit from a stronger focus on parallel activation of prime and target and inspirations from working memory research, working memory research, on the other hand, can be enriched by ideas form research on the influence of similarity that are addressed by priming research. Despite research that focusses on processes that can also be explained by strategic and highly conscious processes induced by relatedness of items like chunking, future research should also focus on automatic processes of similarity on performance in working memory tasks. One way to achieve this might be the use of more subtle forms of similarity. For this purpose, Davelaar and colleagues (2006), as well as Haarman and Usher (2001) and Glanzer (1969), did not use the first associate of a word to generate associated word-pairs. Another strategy, which we had chosen, is the use of evaluative congruency to introduce conceptual similarity. However, using valence as the similarity defining feature could have a disadvantage. Effects might be reduced (or in some tasks even be absent) compared to designs in which stronger semantic associations are used. In priming research, in some tasks, no evaluative priming is observed while semantic priming shows clear effects (Klauer & Musch, 2001, see also the research by Spruyt et al., 2009, 2007, but see M. Becker et al., 2016 for mixed results).

Another idea that can be adopted from priming research is to investigate whether processes work automatically by manipulating the relatedness proportion. In priming, the relatedness proportion is the proportion of trials with related items in the total number of relevant trials (McNamara, 2005). For priming paradigms, it can be assumed that a high relatedness proportion corresponds to a higher prime validity. When the relatedness proportion is high, an expectancy effect can emerge (McNamara, 2005). For example, when a category prime is presented, participants likely generate typical (but not atypical) exemplars of the prime category (Neely, Keefe, & Ross, 1989). Similarly, for list memory tasks, and also the paradigm used by Davelaar and colleagues (2006), the proportion of related word pairs could be reduced to a minimum. Under these circumstances, there should not be the expectation that related word pairs will be presented. Additionally, it can be assumed that conscious generation of associated words, which was sometimes observed in memory tasks (Crowder, 1979; Poirier & Saint-Aubin, 1995), could be less likely when a low relatedness proportion is used. Finding an effect of relatedness when using a low relatedness proportion would provide a hint that the effect of relatedness is not merely based on a conscious generation of related words. If no difference between a high and a low relatedness proportion was observed in these paradigms, this could be regarded as evidence that such conscious processes likely do not influence performance. When there is an effect of the relatedness proportion on the effects of semantic similarity on working memory performance, this could be regarded as evidence for generation processes. It should, however, be noted that these generation processes could either be regarded as strategic or they could be assumed to occur rather automatically (for semantic priming, McNamara, 2005 suggests that the generation processes of semantic sets, which are assumed by C. A. Becker, 1979, are more automatic than strategic).

To differentiate between mutual facilitation in the case of relatedness that boosts performance for related stimuli and a decrease in performance for dissimilar stimuli, priming research provides a valuable tool. That is, neutral stimuli could be used as a baseline and priming research can provide hints as to which stimuli can serve best as a neutral control condition. Therefore, many ideas from priming research can enrich future working memory research that could address effects of semantic relatedness.

If further studies provide evidence for automatic effects of relatedness on working memory performance, as the current Experiment 2a-d and Experiment 4 and the studies by Davelaar and colleagues (2006) (see also Haarmann & Usher 2001) have, the need to theoretically explain these findings will further be endorsed. Besides the theory by Davelaar and colleagues (2006), most working memory theories do not take influences of semantic similarity on performance into account. Based on the evidence for automatic beneficial effects of similarity on working memory performance, the precise mechanisms of these effects have to be further specified. Do these processes occur mainly during encoding, the maintenance of similar concepts, at retrieval or at all three stages? To answer these questions, again, the use of valenced stimuli, as in evaluative priming, seems promising. Stimuli like faces, for which the similarity defining feature (valence) can be separated from the task-relevant feature (the identity of the faces), can be especially useful. Making faces at encoding neutral, and only introducing valence at test (with evaluatively congruent and incongruent test-displays), can show whether there are effects that arise only at test. Presenting evaluatively congruent versus evaluatively incongruent valenced faces at learning, but neutral counterparts at test would reveal whether there are effects of congruency that do *not* emerge at test. An effect in such a task would either occur at encoding or during maintenance. To investigate effects that can only emerge during maintenance, the procedure used by Jackson and colleagues (2014) in Experiment 3 can be used:

Evaluatively congruent displays with only angry faces or only happy faces could be presented at learning. Thereby, congruency cannot have an influence on the encoding of these faces. Presenting a single neutral face as the test item, congruency cannot have an influence on the encoding of the stimulus of the test-display. In their design, congruency or incongruency was added by either an evaluatively congruent or an evaluatively incongruent word that was presented during *maintenance*. Therefore, congruency could not have any influence before maintenance. Jackson and colleagues (2014) observed a significant beneficial effect of evaluative congruency using this design. This effect can be regarded as evidence for a beneficial effect of evaluative congruency on the maintenance of valenced representations (although they interpreted their effect differently). This procedure also creates a link between priming and working memory research. That is because a backward directed influence of similarity is measured just like in the studies investigating retroactive priming (e.g. Dark, 1988; VanVoorhis & Dark, 1995). Therefore, conceptual replications of this study would be of special importance.

6.4.3 Potential assignment processes due to locational changes in change detection

Another kind of implications for working memory research might arise from the difference of results that were obtained depending on whether the stimuli were arranged differently at test compared to encoding or not. In Experiment 3a-c, the beneficial effects of similarity, which were observed in the other working memory studies of this thesis (Experiment 2a-d and Experiment 4), were reversed. One central difference between the Experiments 3a-c and the other current experiments using change detection is that in Experiment 3a-c, the arrangement of the test array differed from the arrangement of the study array. Thereby, in incongruent trials, the emotion of the faces might have been used to assign the stimuli at test to the stimuli from the learning display. More generally, salient features of the test stimuli might make it easier to match test stimuli to the learned stimuli. Therefore, dissimilarity could lead to a performance enhancement if locations of the stimuli change from study to test. As described earlier, this logic also applies to single probed recognition with a centrally present probe. Whether salient features of a single centrally presented probe can function as a post-cue, potentially leading to higher and imprecise measures of working memory capacity, could also be investigated by further research.

6.4.4 Bridging the gap between priming and working memory research

As pointed out above, procedures and methods from working memory research can inspire future priming research and priming research can contribute to working memory research. But also on a theoretical basis, both areas can provide a fruitful foundation for each other. Above, this topic was already touched upon with a focus more on additional research, but here some considerations about linkages between theories are made.

A link between priming and working memory research seems to suggest itself: There is the three process model by Schmitz and Wentura (2012), which states which processes have to be taken into account in priming. One of these processes is the parallel activation of several concepts (namely the prime and the target concept) and parallel activation of several concepts is a core assumption of most working memory models. Furthermore, they assume mutual facilitation of simultaneously active semantically related concepts. This process can explain most of the current results (Experiment 1a-b, Experiment 2a-d and Experiment 4) that originate from both priming and working memory research. Only the third process in their model, the assumed response processes, is arguably more important in priming than in working memory paradigms. Assuming these processes take place in both priming and working memory paradigms, a tremendous amount of findings in both areas can be explained. This includes puzzling findings in recent as well as classical priming studies (see Schmitz & Wentura, 2012), potentially many studies investigating list memory for related words, in which better performance for related words was observed and also most of the results obtained in priming and working memory studies of the current thesis. While the processes that seem to occur in both areas are described in the three process model, the precise mechanisms of how these processes could operate are addressed by other theories.

One model that can convincingly account for findings in both areas and that suggests precise mechanisms for the processes of parallel activation and mutual facilitation is the dual-store neurocomputational model by Davelaar and colleagues (2006). In their model, a spread of activation in a limited capacity short-term buffer (the working memory) is assumed that causes beneficial effects of semantic relatedness in both short-term and long-term memory tasks. Thus, more generally, the application of spreading activation theories, on which the model relies, can account for beneficial effects of semantic similarity. Also, Schmitz and Wentura (2012) suggest that spreading activation theories, but not recent parallel distributed processing models, can account for their evidence for mutual facilitation and parallel activation of related concepts in priming. Thereby, the dual-store neurocomputational model by Davelaar and colleagues (2006), and more generally spreading activation theories, are good candidates to provide a fruitful theoretical foundation to link priming and working memory research. Please note, however, that spreading activation theories like the one by Collins and Loftus (1975) or J. R. Anderson (1976, 1983, 1993) assume localist representations of concepts. Neurologically, however, assuming a parallel distributed processing of semantic concepts seems more plausible, amongst other things because of the robustness of the system and the specific predictions for impairments (e.g. Farah & McClelland, 1991). To give an example, if a part of the system is damaged, models with localist representations would show tremendous impairments. A joke can illustrate this criticism: "Hopefully, after drinking four glasses of wine this evening, I do not lose the one brain cell where my address is stored." Although spreading activation theories might be neurologically implausible, they still have a high potential as a kind of metaphor to easily explain most phenomena in priming research. While parallel distributed processing models are not affected by this criticism, their recent applications to priming research cannot yet account for parallel activation of prime and target. They explain priming by faster transition of the activation pattern of the prime to the activation representing the target, and therefore prime and target are not active concurrently (see Schmitz & Wentura, 2012). New versions of parallel distributed processing models that allow explaining parallel activation of several items and mutual facilitation within the same system (e.g. by implementing interconnected layers to represent several concepts simultaneously) might also provide a promising account to bring working memory and priming research together in a more general framework. Alternatively, as Schmitz and Wentura (2012) suggest, some working memory models that implement synchronous firing for feature nodes that belong to the same concept (e.g Raffone & Van Leeuwen, 2001; Raffone & Wolters, 2001; Vogel et al., 2001; Wolters & Raffone, 2008) could be modified to account for parallel activation of several concepts within parallel distributed processing models. Future research could also address an application of those theories to priming that were originally used to explain effects of perceptual similarity in working memory (Lin & Luck, 2009; Sims et al. 2012).

6.5 Conclusion

In the current thesis, evidence in line with the assumption of parallel activation and mutual facilitation of related concepts in both priming and in working memory is provided. In Experiment 1a and 1b, which were priming studies, evidence for parallel activation and mutual facilitation, for related prime and target concepts, was obtained using a perpetual identification task. Taking the research by Wentura and colleagues (Schmitz & Wentura, 2012; Schmitz et al., 2014; Wentura & Frings, 2008) also into account, two criteria that priming theories should be able to account for are parallel activation of prime and target and mutual facilitation of related prime-target pairs. Thereby, the current data suggest that models of semantic priming should not strongly rely on a sequential activation of prime and target concept. Priming theories that can explain the data and that are compatible with the assumption of a parallel activation are spreading activation theories (e.g. J. R. Anderson, 1976, 1983, 1993) as well as compound-cue theories like the retrieval account by Whittlesea and Jacoby (1990). Further, mutual facilitation can be explained with spreading activation processes. Parallel activation in priming suggests a link between priming and working memory research: Working memory is the system for which it is assumed that several concepts can be maintained simultaneously in an active state. There the mutual facilitation that we assume could take place.

In Experiment 2a-d, evidence in line with the assumption of mutual facilitation of conceptually similar (evaluatively congruent) faces was gathered using a change detection task. Participants performed better when they had to remember four faces expressing the same emotion compared to a condition in which faces expressed different emotions. Not observing this effect with inverted faces, the effect seems not to be merely caused by perceptual overlap; rather, it can be explained by evaluative congruency or more generally by conceptual similarity. The effect in these studies can most parsimoniously be explained by spreading activation theories including the neurocomputational model by Davelaar and colleagues (2006).

In Experiment 3a-c, the beneficial effect of evaluative congruency was not replicated in a similar change detection task, and even reversed. This can be explained by the use of locational changes from study to test in the design. In the incongruent condition, stimuli at test might have been assigned to stimuli of the learning display, based on the distinctive emotion (which was task-irrelevant). This strategy was not applicable in congruent trials. Thereby, this study seems to be the odd one out. Another idea to explain the diverging effects is based on the retrieval account (Whittlesea & Jacoby, 1990) that arguably could jointly explain the findings of the previous studies (when applicable to findings caused by evaluative congruency) as well as the puzzling findings of Experiment 3a-c. The assumption on which this idea is based was empirically tested in Experiment 4.

In Experiment 4, change detection was again used, this time with stimuli presented at the same locations at study and at test as in Experiment 2a-d. This more classical procedure revealed significantly better performance in congruent compared to incongruent trials. This finding again provided evidence for a mutual facilitation of evaluatively congruent concepts in working memory. This last experiment gained similarity to classical priming tasks because only one stimulus was task-relevant at test. The other face (which was evaluatively congruent or evaluatively incongruent) was task-irrelevant. The task-relevant stimulus was marked by a post-cue, enhancing the conceptual similarity to Experiment 1b, in which a somewhat similar procedure was used. In Experiment 4, manipulating whether this task-irrelevant face changed or not allowed us to calculate an interference index. This index was not significantly different in the congruent compared to the incongruent condition. Therefore, no direct evidence for higher unitization in congruent compared to incongruent trials as suggested by the retrieval account by Whittlesea and Jacoby (1990) was obtained. Rather, the results can parsimoniously be explained by the assumption of a mutual facilitation of related concepts as suggested by Schmitz and Wentura (2012) as well as Davelaar and colleagues (2006).

In sum, the current experiments all together imply that linking priming and working memory research is a worthwhile endeavor. The current data, as well as the literature, imply that in both areas, using a wide range of tasks, conceptual similarity can lead to a performance boost. The results of all reported experiments (besides Experiment 3a-c) are in line with the assumption of mutual facilitation of related concepts in working memory (for Experiment 3a-c, other processes seem of importance). The observed beneficial effects of semantic similarity and evaluative congruency could also be explained by a higher unitization of semantically related or evaluatively congruent items, however, in Experiment 4, an attempt to provide direct evidence for this claim failed.

Reversed effects in the experiments with changing locations of the items from study to test (Experiment 3a-c) can be explained by the use of salient features (here the emotion of faces) to assign test stimuli to the studied items. This strategy can only be used in incongruent but not congruent trials. If this strategy did influence results, further implications for working memory research using change detection tasks can be drawn. In single probed recognition with a centrally presented test stimulus, salient but not necessarily task-relevant features could be used as a cue (Wheeler & Treisman, 2002) indicating which item from a study array is tested. This could cause a systematic overestimation of working memory capacity when single probed recognition (with a centrally presented probe) and rather dissimilar stimuli are used. However, further research on this issue is needed.

In the current thesis, another area of working memory research was also addressed: The angry face benefit (Jackson et al., 2008, 2009, 2014). All studies besides

Experiment 1a-b, in which word stimuli were used, allowed for a calculation of this effect. Overall results are in line with the assumption that working memory performance for angry faces is better than for happy faces. However, the effect was not always present. Another interesting finding is that in the analysis of Experiment 2b.2 and Experiment 2c.2 the angry face benefit was also observed for inverted faces. Therefore, in the current studies, the angry face benefit might be partially based on perceptual differences between angry and happy faces.

Most importantly, the major effects of this thesis are mostly in line with the assumption that when several related concepts are in an active state, they mutually facilitate each other's activation.

7 References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*, 106–111. doi:10.1111/j.0963-7214.2004.01502006.x

Anderson, J. E., & Holcomb, P. J. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: An event-related brain potential study. *Psychophysiology*, *32*, 177–190. doi:10.1111/j.1469-8986.1995.tb03310.x

Anderson, J. R. (1976). *Language, memory, and thought*. Oxford, England: Erlbaum.

Anderson, J. R. (1983). *The architecture of cognition*. Hillsdale, NJ, England: Erlbaum.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ, England: Erlbaum.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *The Quarterly Journal of Experimental Psychology*, *18*, 362–365. doi:10.1080/14640746608400055

Baddeley, A. D., & Levy, B. A. (1971). Semantic coding and short-term memory. *Journal of Experimental Psychology*, *89*, 132–136. doi:10.1037/h0031189

Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 336–345. doi:10.1037/0278-7393.12.3.336

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology*, *32*, 104–128. doi:10.1006/jesp.1996.0005

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology*, *25*, 281–316. doi:10.1006/cogp.1993.1007

Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 252–259. doi:10.1037/0096-1523.5.2.252

Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & Cognition*, *8*, 493–512. doi:10.3758/BF03213769

Becker, M., Klauer, K. C., & Spruyt, A. (2016). Is attention enough? A re-examination of the impact of feature-specific attention allocation on semantic priming effects in the pronunciation task. *Attention, Perception, & Psychophysics*, *78*, 396–402. doi:10.3758/s13414-015-1046-7

Berger, T. (1971). *Rate distortion theory: A mathematical basis for data compression*. Englewood Cliffs, NJ: Prentice-Hall.

Bodner, G. E., & Masson, M. E. J. (2003). Beyond spreading activation: An influence of relatedness proportion on masked semantic priming. *Psychonomic Bulletin & Review*, *10*, 645–652. doi:10.3758/BF03196527

Boduroglu, A., & Shah, P. (2009). Effects of spatial configurations on visual change detection: An account of bias changes. *Memory & Cognition*, *37*, 1120–1131. doi:10.3758/MC.37.8.1120

Borowsky, R., & Besner, D. (1993). Visual word recognition: A multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 813–840. doi:10.1037/0278-7393.19.4.813

Bower, G. H. (1991). Mood congruity of social judgments. In J. P. Forgas & J. P. Forgas (Eds.), *Emotion and social judgments.* (pp. 31–53). Elmsford, NY: Pergamon Press.

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, *22*, 384–392. doi:10.1177/0956797610397956

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5):4, 1–34. doi:10.1167/11.5.4

Brady, T. F., Konkle, T., Oliva, A., & Alvarez, G. A. (2009). Detecting changes in real-world objects: The relationship between visual long-term memory and change blindness. *Communicative & Integrative Biology*, *2*, 1–3. doi:10.4161/cib.2.1.7297

Briand, K., den Heyer, K., & Dannenbring, G. L. (1988). Retroactive semantic priming in a lexical decision task. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *40A*, 341–359. doi:10.1080/02724988843000159

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327. doi:10.1111/j.2044-8295.1986.tb02199.x

Calfee, R. C., & Peterson, R. E. (1968). Effect of list organization on short-term probe recall. *Journal of Experimental Psychology*, *78*, 468–474. doi:10.1037/h0026464

Caramazza, A., & Costa, A. (2000). The semantic interference effect in the picture–word interference paradigm: Does the response set matter? *Cognition*, *75*, B51–B64. doi:10.1016/S0010-0277(99)00082-7

Carbon, C.-C., & Leder, H. (2005). When feature information comes first! Early processing of inverted faces. *Perception*, *34*, 1117–1134. doi:10.1068/p5192

Carr, T. H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology. Learning, Memory & Cognition*, *16*, 341–350. doi:10.1037/0278-7393.16.2.341

Center for the Study of Emotion and Attention. (1994). *The International Affective Picture System* [photographic slides]. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.

Chen, Y.-C., & Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology. Human Perception and Performance*, *37*, 1554–1568. doi:10.1037/a0024329

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, M. A., Konkle, T., Rhee, J. Y., Nakayama, K., & Alvarez, G. A. (2014). Processing multiple visual objects is limited by overlap in neural channels. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *111*, 8955–8960. doi:10.1073/pnas.1317860111

Collin, C. A., & McMullen, P. A. (2002). Using Matlab to generate families of similar Attneave shapes. *Behavior Research Methods, Instruments & Computers*, *34*, 55–68. doi:10.3758/BF03195424

Collina, S., Tabossi, P., & Simone, F. D. (2013). Word production and the picture-word interference paradigm: The role of learning. *Journal of Psycholinguistic Research*, *42*, 461–473. doi:10.1007/s10936-012-9229-z

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428. doi:10.1037/0033-295X.82.6.407

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256. doi:10.1037/0033-295X.108.1.204

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*, 163–191. doi:10.1037/0033-2909.104.2.163

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York, NY: Oxford University Press.

Cowan, N. (1999). An Embedded-Processes Model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* (pp. 62–101). New York, NY: Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185. doi:10.1017/S0140525X01003922

Cowan, N. (2005). *Working memory capacity*. New York, NY: Psychology Press. doi:10.4324/9780203342398

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*, 51–57. doi:10.1177/0963721409359277

Cowles, H. W., Garnham, A., & Simner, J. (2010). Conceptual similarity effects on working memory in sentence contexts: Testing a theory of anaphora. *The*

Quarterly Journal of Experimental Psychology, *63*, 1218–1232. doi:10.1080/17470210903359198

Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, *23*, 371–414. doi:10.1016/S0364-0213(99)00005-1

Crowder, R. G. (1979). Similarity and order in memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in research and theory: Vol. 13* (pp. 319–353). Academic Press. doi:10.1016/S0079-7421(08)60086-9

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.

Dagenbach, D., Carr, T. H., & Barnhardt, T. M. (1990). Inhibitory semantic priming of lexical decisions due to failure to retrieve weakly activated codes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 328–340. doi:10.1037/0278-7393.16.2.328

Dale, H. C. A., & Gregory, M. (1966). Evidence of semantic coding in short-term memory. *Psychonomic Science*, *5*, 75–76. doi:10.3758/BF03328287

Dallas, M., & Merikle, P. M. (1976a). Response processes and semantic-context effects. *Bulletin of the Psychonomic Society*, *8*, 441–444. doi:10.3758/BF03335193

Dallas, M., & Merikle, P. M. (1976b). Semantic processing of non-attended visual information. *Canadian Journal of Psychology*, *30*, 15–21. doi:10.1037/h0082040

Dark, V. J. (1988). Semantic priming, prime reportability, and retroactive priming are interdependent. *Memory & Cognition*, *16*, 299–308. doi:10.3758/BF03197040

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review*, *112*, 3–42. doi:10.1037/0033-295X.112.1.3

Davelaar, E. J., Haarmann, H. J., Goshen-Gottstein, Y., & Usher, M. (2006). Semantic similarity dissociates short- from long-term recency effects: Testing a neurocomputational model of list memory. *Memory & Cognition*, *34*, 323–334. doi:10.3758/BF03193410 de Groot, A. M. (1984). Primed lexical decision: Combined effects of the proportion of related prime-target pairs and the stimulus-onset asynchrony of prime and target. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *36A*, 253–280. doi:10.1080/14640748408402158

de Groot, A. M., Thomassen, A. J., & Hudson, P. T. (1982a). Associative facilitation of word recognition as measured from a neutral prime. *Memory & Cognition*, *10*, 358–370. doi:10.3758/BF03202428

de Groot, A. M., Thomassen, A. J., & Hudson, P. T. (1986). Primed-lexical decision: The effect of varying the stimulus-onset asynchrony of prime and target. *Acta Psychologica*, *61*, 17–36. doi:10.1016/0001-6918(86)90019-3

De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition*, *31*, 353–359. doi:10.3758/BF03194393

De Houwer, J., Hermans, D., Rothermund, K., & Wentura, D. (2002). Affective priming of semantic categorisation responses. *Cognition and Emotion*, *16*, 643–666. doi:10.1080/02699930143000419

Dean, M. P., Bub, D. N., & Masson, M. E. J. (2001). Interference from related items in object identification. *Journal of Experimental Psychology. Learning, Memory & Cognition*, *27*, 733-743. doi:10.1037/0278-7393.27.3.733

den Heyer, K., Briand, K., & Smith, L. (1985). Automatic and strategic effects in semantic priming: An examination of Becker's verification model. *Memory & Cognition*, *13*, 228–232. doi:10.3758/BF03197685

Donkin, C., & Nosofsky, R. M. (2012). The structure of short-term memory scanning: An investigation using response time distribution models. *Psychonomic Bulletin & Review*, *19*, 363–394. doi:10.3758/s13423-012-0236-8

Dosher, B. A., & Rosedale, G. (1989). Integrated retrieval cues as a mechanism for priming in retrieval from memory. *Journal of Experimental Psychology: General*, *118*, 191–211. doi:10.1037/0096-3445.118.2.191

Ecker, U. K. H., Maybery, M., & Zimmer, H. D. (2013). Binding of intrinsic and extrinsic features in working memory. *Journal of Experimental Psychology: General*, *142*, 218–234. doi:10.1037/a0028732

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ellamil, M., Susskind, J. M., & Anderson, A. K. (2008). Examinations of identity invariance in facial expression adaptation. *Cognitive, Affective & Behavioral Neuroscience*, *8*, 273–281. doi:10.3758/CABN.8.3.273

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23. doi:10.1111%2F1467-8721.00160

Everaert, T., Spruyt, A., & De Houwer, J. (2011). On the (un)conditionality of automatic attitude activation: The valence proportion effect. *Canadian Journal of Experimental Psychology*, *65*, 125–132. doi:10.1037/a0022316

Everaert, T., Spruyt, A., & De Houwer, J. (2016). Effects in the affect misattribution procedure are modulated by feature-specific attention allocation. *Social Psychology*, *47*, 244–256. doi:10.1027/1864-9335/a000278

Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *33*, 325–350. doi:10.1080/14640748108400797

Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology. General*, *120*, 339–357. doi:10.1037/0096-3445.120.4.339

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. doi:10.3758/BF03193146

Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition*, *25*, 603–637. doi:10.1521/soco.2007.25.5.603

Fischler, I. (1977). Associative facilitation without expectancy in a lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 18–26. doi:10.1037/0096-1523.3.1.18

Fisher, K., Towler, J., & Eimer, M. (2016). Facial identity and facial expression are initially integrated at visual perceptual stages of face processing. *Neuropsychologia*, *80*, 115–125. doi:10.1016/j.neuropsychologia.2015.11.011

Fitousi, D., & Wenger, M. J. (2013). Variants of independence in the perception of facial identity and expression. *Journal of Experimental Psychology. Human Perception and Performance*, *39*, 133–155. doi:10.1037/a0028001

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition and Emotion*, *14*, 61–92. doi:10.1080/026999300378996

Frings, C., Gobel, A., Mast, F., Sutter, J., Bermeitinger, C., & Wentura, D. (2011). Category priming with aliens: Analysing the influence of targets' prototypicality on the centre surround inhibition mechanism. *Memory*, *19*, 585–596. doi:10.1080/09658211.2011.592498

Fuentes, L. J., Carmona, E., Agis, I. F., & Catena, A. (1994). The role of the anterior attention system in semantic processing of both foveal and parafoveal words. *Journal of Cognitive Neuroscience*, *6*, 17–25. doi:10.1162/jocn.1994.6.1.17

Fuentes, L. J., & Tudela, P. (1992). Semantic processing of foveally and parafoveally presented words in a lexical decision task. *The Quarterly Journal of Experimental Psychology Section A*, *45*, 299–322. doi:10.1080/14640749208401328

Gao, T., Gao, Z., Li, J., Sun, Z., & Shen, M. (2011). The perceptual root of object-based storage: An interactive model of perception and visual working memory. *Journal of Experimental Psychology. Human Perception and Performance*, *37*, 1803–1823. doi:10.1037/a0025637

Gao, Z., Flevaris, A. V., Robertson, L. C., & Bentin, S. (2011). Priming global and local processing of composite faces: Revisiting the processing-bias effect on face perception. *Attention, Perception & Psychophysics*, *73*, 1477–1486. doi:10.3758/s13414-011-0109-7

Gao, Z., Li, J., Yin, J., & Shen, M. (2010). Dissociated mechanisms of extracting perceptual information into visual working memory. *PLoS ONE*, *5*(12): e14273, 1–15. doi:10.1371/journal.pone.0014273

Gast, A., Werner, B., Heitmann, C., Spruyt, A., & Rothermund, K. (2014). Evaluative stimulus (in)congruency impacts performance in an unrelated task. Evidence for a resource-based account of evaluative priming. *Experimental Psychology*, *61*, 187–195. doi:10.1027/1618-3169/a000238

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology - Human Perception and Performance*, *10*, 640–654. doi:10.1037/0096-1523.10.5.640

Goh, W. D., & Goh, C. K. (2006). The roles of semantic similarity and proactive interference in the word length effect. *Psychonomic Bulletin & Review*, *13*, 978–984. doi:10.3758/BF03213912

Gupta, P., Lipinski, J., & Aktunc, E. (2005). Reexamining the phonological similarity effect in immediate serial recall: The roles of type of similarity, category cuing, and item recall. *Memory & Cognition*, *33*, 1001–1016. doi:10.3758/BF03193208

Haarmann, H., & Usher, M. (2001). Maintenance of semantic information in capacity-limited item short-term memory. *Psychonomic Bulletin & Review*, *8*, 568–578. doi:10.3758/BF03196193

Hager, W., & Hasselhorn, M. (1994). *Handbuch deutschsprachiger Wortnormen* [Handbook of German word norms]. Göttingen, Germany: Hogrefe.

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of *d'*. *Behavior Research Methods, Instruments, & Computers*, *27*, 46–51. doi:10.3758/BF03203619

Hermans, D., Houwer, J. D., & Eelen, P. (1994). The affective priming effect: Automatic activation of evaluative information in memory. *Cognition and Emotion*, *8*, 515–533. doi:10.1080/02699939408408957

Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., Taylor, J. H., & Crites, S. L. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychological Bulletin*, *139*, 1062–1089. doi:10.1037/a0031309

Hocking, J., McMahon, K. L., & de Zubicaray, G. I. (2010). Semantic interference in object naming: An fMRI study of the postcue naming paradigm. *NeuroImage*, *50*, 796–801. doi:10.1016/j.neuroimage.2009.12.067

Hofer, R. (1965). Intertrial proactive inhibition in short-term memory. *Psychological Reports*, *17*, 755–760. doi:10.2466/pr0.1965.17.3.755

Horstmann, G., & Bauland, A. (2006). Search asymmetries with real faces: Testing the anger-superiority effect. *Emotion*, *6*, 193–207. doi:10.1037/1528-3542.6.2.193

Humphreys, G. W., Lloyd-Jones, T. J., & Fias, W. (1995). Semantic interference effects on naming using a postcue procedure: Tapping the links between semantics and phonology with pictures and words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 961–980. doi:10.1037/0278-7393.21.4.961

Humphreys, M. S., Wiles, J., & Bain, J. D. (1993). Memory retrieval with two cues: Think of intersecting sets. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance 14: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience.* (pp. 489–507). Cambridge, MA: The MIT Press.

Huttenlocher, J., & Newcombe, N. (1976). Semantic effects on ordered recall. *Journal of Verbal Learning and Verbal Behavior*, *15*, 387–400. doi:10.1016/S0022-5371(76)90034-7

Jackson, M. C., Wu, C.-Y., Linden, D. E. J., & Raymond, J. E. (2009). Enhanced visual short-term memory for angry faces. *Journal of Experimental Psychology. Human Perception & Performance*, *35*, 363–374. doi:10.1037/a0013895

Jackson, M. C., Linden, D. E. J., & Raymond, J. E. (2014). Angry expressions strengthen the encoding and maintenance of face identity representations in visual working memory. *Cognition and Emotion*, *28*, 278–297. doi:10.1080/02699931.2013.816655

Jackson, M. C., Linden, D. E. J., Roberts, M. V., Kriegeskorte, N., & Haenschel, C. (2015). Similarity, not complexity, determines visual working memory performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1884–1892. doi:10.1037/xlm0000125

Jackson, M. C., Wolf, C., Johnston, S. J., Raymond, J. E., & Linden, D. E. J. (2008). Neural correlates of enhanced visual short-term memory for angry faces: An fMRI study. *PLoS ONE*, *3*(10): e3536, 1–10. doi:10.1371/journal.pone.0003536

Jackson, M. C., Wu, C.-Y., Linden, D. E. J., & Raymond, J. E. (2009). Enhanced visual short-term memory for angry faces. *Journal of Experimental Psychology. Human Perception & Performance*, *35*, 363–374. doi:10.1037/a0013895

Jiang, Y. V., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 683–702. doi:10.1037/0278-7393.26.3.683

Jiang, Y. V., Lee, H. J., Asaad, A., & Remington, R. (2016). Similarity effects in visual working memory. *Psychonomic Bulletin & Review*, *23*, 476–482. doi:10.3758/s13423-015-0905-5

Jiang, Y. V., Remington, R. W., Asaad, A., Lee, H. J., & Mikkalson, T. C. (2016). Remembering faces and scenes: The mixed-category advantage in visual working memory. *Journal of Experimental Psychology. Human Perception and Performance*, *42*, 1399–1411. doi:10.1037/xhp0000228

Johnson-Laird, P. N., Herrmann, D. J., & Chaffin, R. (1984). Only connections: A critique of semantic networks. *Psychological Bulletin*, *96*, 292–315. doi:10.1037/0033-2909.96.2.292

Klauer, K. C., Becker, M., & Spruyt, A. (2016). Evaluative priming in the pronunciation task: A preregistered replication and extension. *Experimental Psychology*, *63*, 70–78. doi:10.1027/1618-3169/a000286

Klauer, K. C., & Musch, J. (2001). Does sunshine prime loyal? Affective priming in the naming task. *Quarterly Journal of Experimental Psychology: Section A*, *54*, 727–751. doi:10.1080/02724980042000444

Klauer, K. C., & Musch, J. (2013). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion.* (pp. 7–51). Mahwah, NJ: Erlbaum.

Kondo, A., & Saiki, J. (2012). Feature-specific encoding flexibility in visual working memory. *PLoS ONE*, *7*(12): e50962, 1–8. doi:10.1371/journal.pone.0050962

La Heij, W. (1988). Components of stroop-like interference in picture naming. *Memory & Cognition*, *16*, 400–410. doi:10.3758/BF03214220

Lander, K., & Butcher, N. (2015). Independence of face identity and expression processing: Exploring the role of motion. *Frontiers in Psychology*, *6*, 255. doi:10.3389/fpsyg.2015.00255

Langeslag, S. J. E., Morgan, H. M., Jackson, M. C., Linden, D. E. J., & Van Strien, J. W. (2009). Electrophysiological correlates of improved short-term memory for emotional faces. *Neuropsychologia*, *47*, 887–896. doi:10.1016/j.neuropsychologia.2008.12.024

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology. Human Perception and Performance*, *21*, 451–468. doi:10.1037/0096-1523.21.3.451

Liberman, N., & Förster, J. (2009). Distancing from experienced self: How global-versus-local perception affects estimation of psychological distance. *Journal of Personality & Social Psychology*, *97*, 203–216. doi:10.1037/a0015671

Lin, P.-H., & Luck, S. J. (2009). The influence of similarity on visual working memory representations. *Visual Cognition*, *17*, 356–372. doi:10.1080/13506280701766313

Liu, C. C., & Watanabe, T. (2012). Accounting for speed–accuracy tradeoff in perceptual learning. *Vision Research*, *61*, 107–114. doi:10.1016/j.visres.2011.09.007

Loess, H. (1964). Proactive inhibition in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *3*, 362–368. doi:10.1016/S0022-5371(64)80003-7

Loess, H. (1967). Short-term memory, word class, and sequence of items. *Journal of Experimental Psychology*, *74*, 556–561. doi:10.1037/h0024787

Loess, H. (1968). Short-term memory and item similarity. *Journal of Verbal Learning & Verbal Behavior*, *7*, 87–92. doi:10.1016/S0022-5371(68)80169-0

Logie, R. H., Brockmole, J. R., & Jaswal, S. (2011). Feature binding in visual short-term memory is unaffected by task-irrelevant changes of location, shape, and color. *Memory & Cognition*, *39*, 24–36. doi:10.3758/s13421-010-0001-z

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review*, *7*, 618–630. doi:10.3758/BF03212999

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281. doi:10.1038/36846

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*, 391–400. doi:10.1016/j.tics.2013.06.006

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces (KDEF).* Stockholm: Department of Neurosciences Karolinska Hospital.

Mahr, A., & Wentura, D. (2014). Time-compressed spoken word primes crossmodally enhance processing of semantically congruent visual targets. *Attention, Perception, & Psychophysics*, *76*, 575–590. doi:10.3758/s13414-013-0569-z

Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 369–380. doi:10.1037/0278-7393.34.2.369

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*, 1–22. doi:10.1037/0096-3445.111.1.1

Masson, M. E. J. (1991). A distributed memory model of context effects in word identification. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition.* (pp. 233–263). Hillsdale, NJ, England: Erlbaum.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 3–23. doi:10.1037/0278-7393.21.1.3

Masson, M. E. J., & Bodner, G. E. (2003). A retrospective view of masked priming: Toward a unified account of masked and long-term repetition priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art.* (pp. 57–94). New York, NY: Psychology Press.

MATLAB and Statistics Toolbox Release 2016a, The MathWorks, Inc., Natick, Massachusetts, United States.

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

McKelvie, S. J. (1995). Emotional expression in upside-down faces: Evidence for configurational and componential processing. *British Journal of Social Psychology*, *34*, 325–334. doi:10.1111/j.2044-8309.1995.tb01067.x

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York, NY: Psychology Press.

McNamara, T. P. (2013). Semantic memory and priming. In A. F. Healy, R. W. Proctor, & I. B. Weiner (Eds.), *Handbook of psychology, Vol. 4: Experimental psychology* (2nd ed., pp. 449–471). Hoboken, NJ: Wiley.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*, 99–130. doi:10.1037/0096-3445.126.2.99

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*, 227–234. doi:10.1037/h0031564

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97. doi:10.1037/h0043158

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*, 297–326. doi:10.1037/0033-2909.132.2.297

Moss, H. E., Hare, M. L., Day, P., & Tyler, L. K. (1994). A distributed memory model of the associative boost in semantic priming. *Connection Science*, *6*, 413–427. doi:10.1080/09540099408915732

Murdock, B. B. J., & Vom Saal, W. (1967). Transpositions in short-term memory. *Journal of Experimental Psychology*, *74*, 137–143. doi:10.1037/h0024507

Murphy, K. (2010). Semantic priming occurs for word but not location pronunciation in the postcue task. *Cognitive Processing*, *12*, 197–201. doi:10.1007/s10339-010-0381-9

Murphy, K., & Green, L. (2011). The effect of target context and cue type in a postcue word pronunciation task. *Advances in Cognitive Psychology*, *7*, 39–48. doi:10.2478/v10053-008-0086-0

Mutluturk, A., & Boduroglu, A. (2014). Effects of spatial configurations on the resolution of spatial working memory. *Attention, Perception, & Psychophysics*, *76*, 2276–2285. doi:10.3758/s13414-014-0713-4

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254. doi:10.1037/0096-3445.106.3.226

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition.* (pp. 264–336). Hillsdale, NJ: Erlbaum.

Neely, J. H., & Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective-retrospective processing theory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 24, pp. 207–248). New York: Academic Press.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1003–1019. doi:10.1037/0278-7393.15.6.1003

Nisbett, R. E., & Miyamoto, Y. (2005). The influence of culture: Holistic versus analytic perception. *Trends in Cognitive Sciences*, *9*, 467–473. doi:10.1016/j.tics.2005.08.004

Nosofsky, R. M., & Donkin, C. (2016). Response-time evidence for mixed memory states in a sequential-presentation change-detection task. *Cognitive Psychology*, *84*, 31–62. doi:10.1016/j.cogpsych.2015.11.001

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*, 411–421. doi:10.1037//0278-7393.28.3.411

Oberauer, K. (2006). Is the focus of attention in working memory expanded through practice? *Journal of Experimental Psychology: Learning Memory and Cognition*, *32*, 197–214. doi:10.1037/0278-7393.32.2.197

Oberauer, K. (2009a). Design for a working memory. In B. H. Ross (Ed.), *Psychology of learning and motivation: Advances in research and theory* (Vol. 51, pp. 45–100). San Diego, CA: Elsevier Academic Press.

Oberauer, K. (2009b). Interference between storage and processing in working memory: Feature overwriting, not similarity-based competition. *Memory & Cognition*, *37*, 346–357. doi:10.3758/MC.37.3.346

Oberauer, K., & Lange, E. B. (2009). Activation and binding in verbal working memory: A dual-process model for the recognition of nonwords. *Cognitive Psychology*, *58*, 102–136. doi:10.1016/j.cogpsych.2008.05.003

Oberauer, K., Souza, A. S., Druey, M. D., & Gade, M. (2013). Analogous mechanisms of selection and updating in declarative and procedural working memory: Experiments and a computational model. *Cognitive Psychology*, *66*, 157–211. doi:10.1016/j.cogpsych.2012.11.001

Olivers, C. N. L., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology. Human Perception and Performance*, *32*, 1243–1265. doi:10.1037/0096-1523.32.5.1243

Olsson, H., & Poom, L. (2005). Visual memory needs categories. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 8776–8780. doi:10.1073/pnas.0500810102

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*, 369–378. doi:10.3758/BF03210419

Pearson, B., Raškevičius, J., Bays, P. M., Pertzov, Y., & Husain, M. (2014). Working memory retrieval as a decision process. *Journal of Vision*, *14*(2):2, 1–15. doi:10.1167/14.2.2

Pecher, D., Zeelenberg, R., & Barsalou, L. w. (2003). Verifying different–modality properties for concepts produces switching costs. *Psychological Science*, *14*, 119–124. doi:10.1111/1467-9280.t01-1-01429

Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. W. (2002). Associative priming in a masked perceptual identification task: Evidence for automatic processes. *The Quarterly Journal of Experimental Psychology: Section A*, *55*, 1157–1173. doi:10.1080/02724980244000143

Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. In *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 37-42). Hillsdale, NJ: Erlbaum.

Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review*, *107*, 786–823. doi:10.1037/0033-295X.107.4.786

Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words: Further evidence on the influence of semantic factors in immediate serial recall. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *48*, 384–404. doi:10.1080/14640749508401396

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 785–794. doi:10.1037/0278-7393.21.3.785

Prkachin, G. C. (2003). The effects of orientation on detection and identification of facial expressions of emotion. *British Journal of Psychology*, *94*, 45. doi:10.1348/000712603762842093

Raffone, A., & Van Leeuwen, C. (2001). Activation and coherence in memory processes: Revisiting the parallel distributed processing approach to retrieval. *Connection Science*, *13*, 349–382. doi:10.1080/09540090110109173

Raffone, A., & Van Leeuwen, C. (2003). Dynamic synchronization and chaos in an associative neural network with multiple active memories. *Chaos*, *13*, 1090-1104. doi:10.1063/1.1602111

Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience*, *13*, 766–785. doi:10.1162/08989290152541430

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108. doi:10.1037/0033-295X.85.2.59

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, *111*, 159–182. doi: 10.1037/0033-295X.111.1.159

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*, 385–408. doi:10.1037/0033-295X.95.3.385

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922. doi:10.1162/neco.2008.12-06-420

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*, 347–356. doi:10.1111/1467-9280.00067

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*, 333–367. doi:10.1037/0033-295X.111.2.333

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, *47*, 25–57. doi:10.1016/0010-0277(93)90061-Y

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, *12*, 1–20. doi:10.1016/S0022-5371(73)80056-8

Rouder, J., Morey, R., Morey, C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*, *18*, 324–330. doi:10.3758/s13423-011-0055-3

Russell, W. A. (1970). The complete German language norms for responses to 100 words from the Kent-Rosanoff word association test. In L. Postman & G. Keppel (Eds.), *Norms of Word Association* (pp. 53–94). New York: Academic Press.

Saint-Aubin, J., Ouellette, D., & Poirier, M. (2005). Semantic similarity and immediate serial recall: Is there an effect on all trials? *Psychonomic Bulletin & Review*, *12*, 171–177. doi:10.3758/BF03196364

Scherer, D. (2015). *Semantisches Priming in der perzeptuellen Identifikationsaufgabe. Eine Vorstudie zur Messung der parallelen Aktivierung von Prime und Target* (Unpublished master's thesis). Saarland University, Saarbrücken, Germany.

Scherer, D., & Wentura, D. (2018). Combining the post-cue task and the perceptual identification task to assess parallel activation and mutual facilitation of related primes and targets. *Experimental Psychology*, *65*, 84–97. doi:10.1027/1618-3169/a000396

Schmitz, M. (2012). *Evaluative priming in non-evaluative priming tasks: A mutual facilitation account* (Doctoral dissertation). Saarland University, Saarbrücken, Germany. doi:10.22028/D291-23351

Schmitz, M., & Wentura, D. (2012). Evaluative priming of naming and semantic categorization responses revisited: A mutual facilitation explanation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 984–1000. doi:10.1037/a0026779

Schmitz, M., Wentura, D., & Brinkmann, T. A. (2014). Evaluative priming in a semantic flanker task: ERP evidence for a mutual facilitation explanation. *Cognitive, Affective & Behavioral Neuroscience*, *14*, 426–442. doi:10.3758/s13415-013-0206-2

Schneider, W., Eschman, A., & Zuccolotto, A. P. (2002). *E-Prime user's guide.* Pittsburgh, PA: Psychology Software Tools, Inc.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, *29*, 86–102. doi:10.1016/0749-596X(90)90011-N

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial–relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 904–915. doi:10.1037/0096-1523.22.4.904

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory & Cognition*, *12*, 315–328. doi:10.3758/BF03198291

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Champaign, IL: University of Illinois Press.

Sharkey, A. J., & Sharkey, N. E. (1992). Weak contextual constraints in text and word priming. *Journal of Memory and Language*, *31*, 543–572. doi:10.1016/0749-596X(92)90028-V

Shivde, G., & Anderson, M. C. (2011). On the existence of semantic working memory: Evidence for direct semantic maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1342–1370. doi:10.1037/a0024832

Sims, C. R., Jacobs, R. A., & Knill, D. C. (2012). An ideal observer analysis of visual working memory. *Psychological Review*, *119*, 807–830. doi:10.1037/a0029856

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*, 214–241. doi:10.1037/h0036351

Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology. Learning, Memory & Cognition*, *32*, 101–117. doi:10.1037/0278-7393.32.1.101

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*, 1–29. doi:10.1037/h0093759

Spruyt, A., De Houwer, J., & Hermans, D. (2009). Modulation of automatic semantic priming by feature-specific attention allocation. *Journal of Memory and Language*, *61*, 37–54. doi:10.1016/j.jml.2009.03.004

Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experimental Psychology*, *54*, 44–53. doi:10.1027/1618-3169.54.1.44

Spruyt, A., Hermans, D., Houwer, J. D., & Eelen, P. (2002). On the nature of the affective priming effect: Affective priming of naming responses. *Social Cognition*, *20*, 227–256. doi:10.1521/soco.20.3.227.21106

Spruyt, A., Hermans, D., Pandelaere, M., De Houwer, J., & Eelen, P. (2004). On the replicability of the affective priming effect in the pronunciation task. *Experimental Psychology*, *51*, 109–115. doi:10.1027/1618-3169.51.2.109

Thomas, P. M. J., Jackson, M. C., & Raymond, J. E. (2014). A threatening face in the crowd: Effects of emotional singletons on visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 253–263. doi:10.1037/a0033970

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

Turvey, M. T., Cremins, J. J., & Lombardo, T. (1969). Taxonomic categories and proactive interference in short-term memory. *Psychonomic Science*, *15*, 307–308. doi:10.3758/BF03336313

Usher, M., & Cohen, J. D. (1999). Short term memory and selection processes in a frontal-lobe model. In D. Heinke, G. W. Humphreys, & A. Olson (Eds.), *Connectionist Models in Cognitive Neuroscience* (pp. 78–91). Berlin, Germany: Springer. doi:10.1007/978-1-4471-0813-9_7

Van der Lely, H. K. J., & Howard, D. (1993). Children with specific language impairment: Linguistic impairment or short-term memory deficit? *Journal of Speech and Hearing Research*, *36*, 1193–1207. doi:10.1044/jshr.3606.1193

Van Veen, V., & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, *14*, 593–602. doi:10.1162/08989290260045837

Van Vugt, M. K., & Jha, A. P. (2011). Investigating the impact of mindfulness meditation training on working memory: A mathematical modeling approach. *Cognitive, Affective & Behavioral Neuroscience*, *11*, 344–353. doi:10.3758/s13415-011-0048-8

VanVoorhis, B. A., & Dark, V. J. (1995). Semantic matching, response mode, and response mapping as contributors to retroactive and proactive priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 913–932. doi:10.1037/0278-7393.21.4.913

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology:*

Human Perception and Performance, *27*, 92–114. doi:10.1037/0096-1523.27.1.92

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology: A practical introduction. *Experimental Psychology*, *60*, 385–402. doi:10.1027/1618-3169/a000218

Voss, A., Rothermund, K., Gast, A., & Wentura, D. (2013). Cognitive processes in associative and categorical priming: A diffusion model analysis. *Journal of Experimental Psychology. General*, *142*, 536–559. doi:10.1037/a0029459

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*, 1206–1220. doi:10.3758/BF03196893

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, *39*, 767–775.

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology*, *52*, 1–9. doi:10.1016/j.jmp.2007.09.005

Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: A tutorial based on fast-dm-30. *Frontiers in Psychology*, *6*. doi:10.3389/fpsyg.2015.00336

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, *21*, 641–671. doi:10.1080/09541440802205067

Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*, 3–22. doi:10.3758/BF03194023

Warrington, E. K., Kinsbourne, M., & James, M. (1966). Uncertainty and transitional probability in the span of apprehension. *British Journal of Psychology*, *57*, 7–16. doi:10.1111/j.2044-8295.1966.tb00998.x

Wentura, D., & Degner, J. (2010). A practical guide to sequential priming and related tasks. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social* *cognition: Measurement, theory, and applications.* (pp. 95–116). New York, NY: Guilford Press.

Wentura, D., & Frings, C. (2008). Response-bound primes diminish affective priming in the naming task. *Cognition and Emotion*, *22*, 374–384. doi:10.1080/02699930701446064

Wentura, D., & Rothermund, K. (2003). The “meddling-in” of affective information: A general model of automatic evaluation effects. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion.* (pp. 51–86). Mahwah, NJ: Erlbaum.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology. General*, *131*, 48–64. doi:10.1037/0096-3445.131.1.48

Whittlesea, B. W. A., & Jacoby, L. L. (1990). Interaction of prime repetition with visual degradation: Is priming a retrieval phenomenon? *Journal of Memory and Language*, *29*, 546–565. doi:10.1016/0749-596X(90)90051-Z

Wickens, D. D. (1970). Encoding categories of words: An empirical approach to meaning. *Psychological Review*, *77*, 1–15. doi:10.1037/h0028569

Wickens, D. D., & Clark, S. (1968). Osgood dimensions as an encoding class in short-term memory. *Journal of Experimental Psychology*, *78*, 580–584. doi:10.1037/h0026643

Wickens, D. D., Moody, M. J., & Dow, R. (1981). The nature and timing of the retrieval process and of interference effects. *Journal of Experimental Psychology: General*, *110*, 1–20. doi:10.1037/0096-3445.110.1.1

Wilford, M. M., & Wells, G. L. (2010). Does facial processing prioritize change detection? Change blindness illustrates costs and benefits of holistic processing. *Psychological Science*, *21*, 1611–1615. doi:10.1177/0956797610385952

Wolters, G., & Raffone, A. (2008). Coherence and recurrency: Maintenance, control and integration in working memory. *Cognitive Processing*, *9*, 1–17. doi:10.1007/s10339-007-0185-8

Woodman, G. F., Vogel, E. K., & Luck, S. J. (2012). Flexibility in visual working memory: Accurate change detection in the face of irrelevant variations in position. *Visual Cognition*, *20*, 1–28. doi:10.1080/13506285.2011.630694

Yu, J.-C., Chang, T.-Y., & Yang, C.-T. (2014). Individual differences in working memory capacity and workload capacity. *Frontiers in Psychology*, *5,* 1465. doi:10.3389/fpsyg.2014.01465

Zhang, J., & Rowe, J. B. (2014). Dissociable mechanisms of speed-accuracy tradeoff during visual perceptual learning are revealed by a hierarchical drift-diffusion model. *Frontiers in Neuroscience*, *8*, 69. doi:10.3389/fnins.2014.00069

Zhou, J., Yin, J., Chen, T., Ding, X., Gao, Z., & Shen, M. (2011). Visual working memory capacity does not modulate the feature-based information filtering in visual working memory. *PloS One*, *6*(9): e23873, 1–10. doi:10.1371/journal.pone.0023873

Zimmer, H., Mecklinger, A., & Lindenberger, U. (2006). *Handbook of binding and memory: Perspectives from cognitive neuroscience*. Oxford University Press.

8 Appendix

8.1 Appendix A

Supporting Analysis: Analysis of carry-over effects (Experiment 2b)

As outlined in the Introduction of Experiment 2a-d, Experiment 2b originally had a counter-balanced design with orientation (upright vs. inverted) as a blocked within-participants factor. Thus, for the sake of completeness and transparency we present here the corresponding 2 (congruency: congruent vs. incongruent) × 2 (orientation: upright vs. inverted) × 2 (order: upright faces first vs. inverted faces first) ANOVA with the first two factors as within-participants factors and the last factor as a between-participants factor (see Table 13 for the descriptive statistics).

Table 13 *Mean performance (in d') for upright and inverted faces as a function of congruency depending on the order of blocks in Experiment 2b*

	Sample Upright first		Sample Inverted first	
	Mean Congruent (*SD*)	Mean Incongruent (*SD*)	Mean Congruent (*SD*)	Mean Incongruent (*SD*)
Upright	1.71 (0.61)	1.57 (0.60)	1.28 (0.58)	1.44 (0.53)
Inverted	1.34 (0.63)	1.23 (0.60)	1.07 (0.43)	1.07 (0.54)
Total	1.52 (0.51)	1.40 (0.50)	1.17 (0.44)	1.26 (0.46)

For this analysis, one participant, who did not reach a performance above chance in the condition with inverted faces, was excluded. Besides a main effect of inversion ($F(1, 65) = 24.32$, $p < .001$, $\eta_p^2 = .272$), there was only a significant interaction of congruency and order, $F(1,65) = 7.47$, $p = .008$, $\eta_p^2 = .103$ – all other $p > .193$. This significant interaction indicates that the effect of congruency depended on the factor order that indicates which face stimuli were presented first – the upright or the inverted ones. Participants that firstly performed the task on upright faces showed overall (i.e., collapsed for upright and inverted faces) better performance in congruent ($M = 1.52$, $SD = 0.51$) compared to incongruent trials ($M = 1.40$, $SD = 0.50$), $t(32) = 2.22$, $p = .034$, $d_Z = .39$.) Participants that first performed the task with inverted faces showed the opposite pattern. They were numerically better in incongruent ($M = 1.26$, $SD = 0.46$) compared to congruent trials ($M = 1.17$, $SD = 0.44$), $t(33) = 1.62$, $p = .115$, $d_Z = .28$. These analyses provide evidence for carry-over effects that might have interesting implications (that we touch upon in the discussion of the experiment, in section 3.5.1, and the discussion at the end of the thesis, in section 6.2.8); however, for the analysis of Experiment 2a-d, they indicate that only the first block of Experiment 2b is comparable to the methods used in the other experiments. Thus, only the first block entered the overall analysis of Experiment 2a-d.

8.2 Appendix B

Detailed Methods of Experiment 3a-c

B1 Experiment 3a

As noted in the description of Experiment 3a-c, the main aim of Experiment 3a was to investigate the effect of evaluative congruency on working memory performance using a change detection task. Despite the fact that participants were required to give several responses within a trial (see below), only performance in the central change detection task for face identities, but no other responses, entered the analyses.

B1.1 Methods

Participants. 65 students (46 females, age range: 18-32, *Md* = 24 years) took part in Experiment 3a and were paid 10 € for participation. All of them had normal or corrected-to-normal vision. The data of 4 further participants were excluded from the analysis because they did not perform above chance in the identity change detection task. Assuming a small to medium effect size of $d_Z = .30$ (that is above the one observed in Experiment 1), testing $N = 70$ participants would lead to a power of $1\text{-}\beta = .80$.

Design and procedure. Participants performed a change detection task in which evaluative congruency was manipulated, which followed the procedure described in the methods section of Experiment 3a-c. In this experiment, the task was made more difficult by inducing a higher load. This was achieved by letting participants remember two abstract shapes that were presented prior to the learning display with two faces. After the response to the test display containing two faces, a second test display was presented, showing again two abstracts shapes. These two shapes were either the same as at the beginning of the trial or one shape was replaced by another one. By presenting the test display with the faces before the test display with the two shapes, it is ensured that responses in the critical change detection task using faces are not biased by a previous response to the shapes. This method resulted in a sandwich-like trial procedure with a change detection task for faces in the middle of each trial, surrounded by a change detection task for shapes. In both tasks, the same two keys were used to indicate whether there was a match or a change. In contrast to the faces that changed locations from study to test, the positions of the shapes were the same at learning and at test.

In addition to these trials, trials with a different task for the faces were incorporated into the design. In one-third of the trials, after the learning display containing two faces, there was no test display with faces. Instead, prior to the test display containing shapes, there was a different test display in which a single location of a previously seen face was marked by a frame. In this frame "pos" as an abbreviation for positive or "neg" for negative was written. In these trials, participants had to indicate whether the face, which was previously presented at this position, had the indicated valence or not. If there was a match of valence, participants pressed the key for matches, which was also used in the change detection tasks for the face identities and the change detection task for shapes (e.g. if the left position was marked by a frame in which "pos" was written, and the face presented previously on the left side was positively valenced, they had to indicate via keypress that there was no change). If there was a change of valence, they had to press the key for a change (e.g. when the right position was marked with a frame in which "neg" was written, but the face on the right side expressed happiness, indicating a change is the correct response). During study and during maintenance of the two faces, participants did not have any information about which test display would appear. Therefore, prior to the presentation of the test display, they did not know whether memory for valence or face identity would be tested.

The procedure of a single trial is illustrated in Figure 14. Each trial started with a sequence of a plus sign, followed by an "X" and again by a plus sign, all presented in the middle of the screen. This sequence indicated the start of a trial. After that, participants saw two random shapes, which they had to remember, followed by a short blank and then the two to-be-remembered faces. During maintenance, a blank was presented. Then in one-third of the trials the emotion was tested and participants had to indicate whether the expression in the marked position ("pos'" for positive, "neg" for negative) matched the valence of the face previously presented at this position. In two third of the trials, a test display containing two faces with one face presented above the other was shown. Participants indicated via key press whether these two faces were the same as before or whether one face was replaced by another one. After a response was given in either of both test display types, a blank was presented (for 1000 ms) if the response was right. Otherwise, the word “Fehler” (error) was presented for the same duration (1000 ms, not depicted in Figure 14). After that, a further test display containing two shapes was presented. Again, participants indicated whether one stimulus was replaced by another one or not using the same keys as before. Also after their response to the shapes, either a blank or an error feedback showing “Fehler” (error) was presented for 1000 ms (not shown in Figure 14) before the next trial started.

Furthermore, after each block of 48 trials (and after each part of the practice phase, see below) feedback was presented showing the percentage correct of all three parts of the task: Participants saw their performance for the identity-change detection task, the valence change detection task, and the shape-change detection task, all three indicated in percentage correct. In addition, participants that performed in either of these subtasks below 60% were encouraged to engage more in this aspect of the procedure. If participants were more than 5 % better in the identity change detection task than in the shape change detection task, they were encouraged to focus more on the shape change detection task.

The experiment started with an extensive practice phase, in which the task was introduced to the participants. They first completed 16 practice trials with the identity change detection task surrounded by the change detection task for shapes. Participants were instructed to try to achieve a comparable performance in both tasks, the change detection task for the identities and the change detection task for the shapes. In these first practice trials, there were no trials in which valence was task-relevant. Following this first practice, there was a second practice phase with 16 trials in which only the memory for valence was tested (surrounded by the change detection task for the shapes). Following this, participants were informed that there would be both types of trials. They were informed that, in two third of trials, they would see two faces at test and that they have to indicate whether these two faces were the same as before or whether one identity changed. They were further informed that in one-third of trials, a position will be marked, and they have to decide whether the term "pos" or "neg" matches the valence of the previously seen face that was presented on the marked position. Then there was a third final practice phase with 32 trials in which the memory for the identities was tested and 16 trials in which the memory for the valence of a single face was tested (again, both kinds of trials were surrounded by the change detection task for the shapes). As in the main phase, these different trials were randomly intermixed.

After participants got familiar with the tasks, the main phase of the experiment started. In this phase again, as previously described, in one-third of the trials the emotion was probed. In two third of the trials, the change detection task with a test for the identity of the faces was utilized. These trials were relevant for the test of our hypotheses. The main phase consisted of 4 blocks with 48 trials each (32 relevant identity change detection trials and 16 valence trials in each block). Each of both change detection tasks (for identities and for valence) was surrounded by the change detection task for shapes.

The size of the face stimuli on the screen and other procedural details are described in depth in the common design and procedure section (4.1.1). The shapes that were only used in Experiment 3a were presented almost next to each other in the middle of the screen (both at study and at test), covering together a visual angle of up to $6.62 \times 3.31°$ visual angle (a single shape could cover up to $3.31° \times 3.31°$).

Materials. There were two sets of face stimuli (Set A and Set B) each containing 9 different identities (see overview). Half of the participants saw happy faces from Set A and angry faces from Set B. Due to a programming error, however, the other half of participants saw angry faces as well as happy faces from Set A. In addition, for these participants it was possible that the same identity was presented twice within a trial expressing different emotions. These trials were discarded (a total of 128 trials, 0-9 trials per participant) and results stayed the same no matter whether they were included in the analysis or not. Furthermore, results were highly comparable for participants only seeing faces from Set A and participants for which happy faces from Set A and angry faces from Set B were presented. In addition, there is no obvious argument as to why this could lead to different findings for the congruency effect.

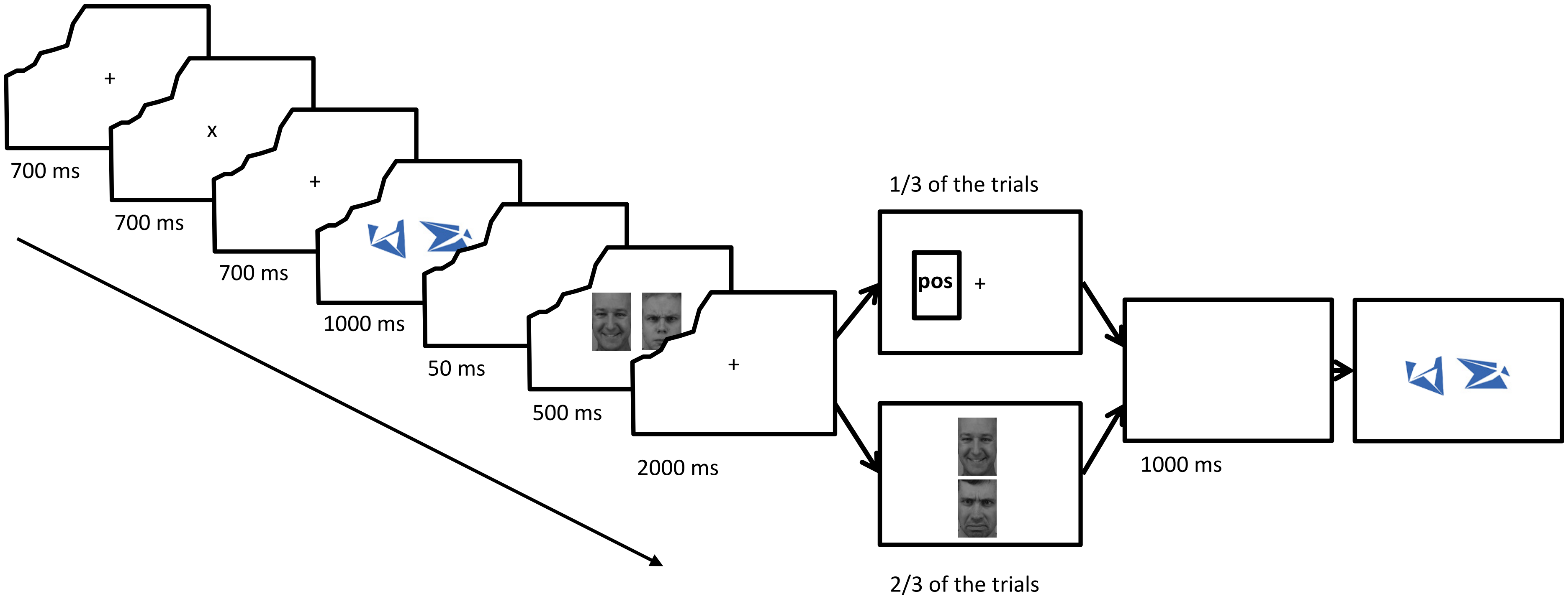


Figure 14. *Trial sequence of the change detection task in Experiment 3a. (Depicted is a change trial with respect to the identity and a trial that requires a no change response regarding the emotion).*

Additionally, a set of 9 random shapes (with no common prototype) was utilized, generated using Matlab (The Mathworks, Inc., Natick, MA; www.mathworks.com) and a script by Collin and McMullen (2002). These shapes were colored blue (RGB: 55,103,176) using Adobe Photoshop. Shapes with 12 to 19 edges and comparable luminance of the total picture (the shape plus the white background) were selected.

B1.2 Results and Discussion

The results of Experiment 3a-c are reported in the common results section. There, a tendency for better performance in incongruent compared to congruent trials is reported for Experiment 3a. The analysis of response times revealed faster responses in congruent compared to incongruent trials (an effect that was significantly moderated by the factor change type, but the effect was significant in change trials as well as no change trials). Therefore, analysis of *d'* and analysis of reaction times revealed a tendency for diverging effects. Due to the insignificant effect for *d'*, reaction times might be more important; however, in contrast to paradigms or tasks in which reaction times are the main focus of analysis, in our paradigm accuracy was rather low.

In a further experiment we could investigate whether the effect in reaction times and the tendency in accuracy can be replicated. However, a direct replication could again provide hints for a speed-accuracy trade-off. Thus, an alternative could be a replication in which there are more correct trials, to enable a more meaningful analysis of reaction time data. Otherwise much less correct trials could transmit any effects in reaction times to effects in accuracy. The latter could be achieved by further enhancing task difficulty. Alternatively, the former can be achieved by making the task easier and the current design less complex (e.g. by leaving out the shape-change detection surrounding the main task) to enable a more constructive analysis of reaction time data. A third way to gain more insight into the nature of the observed pattern of results in this task would be a combined analysis of both measurements. One way to realize this is drift-diffusion modeling (Ratcliff & Rouder, 1998; Voss, Nagler, et al., 2013; Wagenmakers, Maas, & Grasman, 2007). However, directly performing a drift-diffusion analysis of the current data would be inappropriate, because, with data sets of small or medium size, models would not be stable especially when independent models are estimated for several conditions (i.e., the congruent and the incongruent condition). Actually, we had a total of 128 relevant trials in this experiment, and Voss and colleagues (2013) state that having under 200 trials might lead to unstable parameter estimations. It can be argued that larger sample sizes can compensate for a low amount of trials; however, for that purpose, the current sample alone is not sufficient.

B2 Experiment 3b

In Experiment 3b, the effect of evaluative congruency on working memory performance was again investigated, employing a change detection task with set size two and locational changes from study to test. Experiment 3b was designed in a way to obtain a higher performance (and to allow for a more informative analysis of reaction times). Therefore, the additional change detection for shapes (see Experiment 3a) was dropped, the presentation time of the study array was prolonged and the maintenance period was shortened. The reduced load, the longer presentation time of the study array (Brady, Konkle, Oliva, & Alvarez, 2009; Mandler & Shebo, 1982; Vogel et al., 2001) as well as the shorter lag between study and test display (Donkin & Nosofsky, 2012; Nosofsky & Donkin, 2016; Pashler, 1988 but see Logie et al., 2011) all are factors that typically lead to a higher performance in change detection paradigms (for an extensive discussion of factors affecting working memory performance see also Brady et al., 2011; Cowan, 2005). All other features of the task remained virtually the same.[26]

B2.1 Methods

70 students (56 females, age range: 18-34, *Md* = 23 years) were paid 6 € for participation. All of them reported normal or corrected-to-normal vision. Assuming an effect size as the one of the congruency effect in reaction times in Experiment 3a of $d_Z = .61$, with $N = 70$ and $\alpha = .05$ (two-tailed), a power of $1\text{-}\beta = 1$ (.9997) would be obtained. Alternatively, an effect of the same size as the tendency for an incongruency effect in the accuracy (measured with *d’*) in Experiment 3b ($d_Z = .20$) would be observed with a power of $1\text{-}\beta = .50$; a medium effect of $d_Z = .35$ would be detected with $1\text{-}\beta = .90$.

Design and Procedure. The design and the procedure were virtually the same as in Experiment 3a; however, participants did not have to remember shapes prior to the presentation of the study array for the faces. Again, in one-third of the trials, a different test display was presented, probing participant’s memory for

[26] The programming error from Experiment 3a was fixed.

valence. As in Experiment 3a, 128 relevant trials (from 4 blocks, each including 32 relevant identity change detection trials) of the main phase entered the analysis. Note, however, that the experiment was shortened in its duration by the changed trial sequence without the change detection task for shapes surrounding the change detection task for face identities (or surrounding the change detection task for valence, in one-third of the trials) as in Experiment 3a. The trial sequence of Experiment 3b, with the longer presentation time for the study array and the shortened presentation time for the blank during maintenance, is presented in

Figure 15.

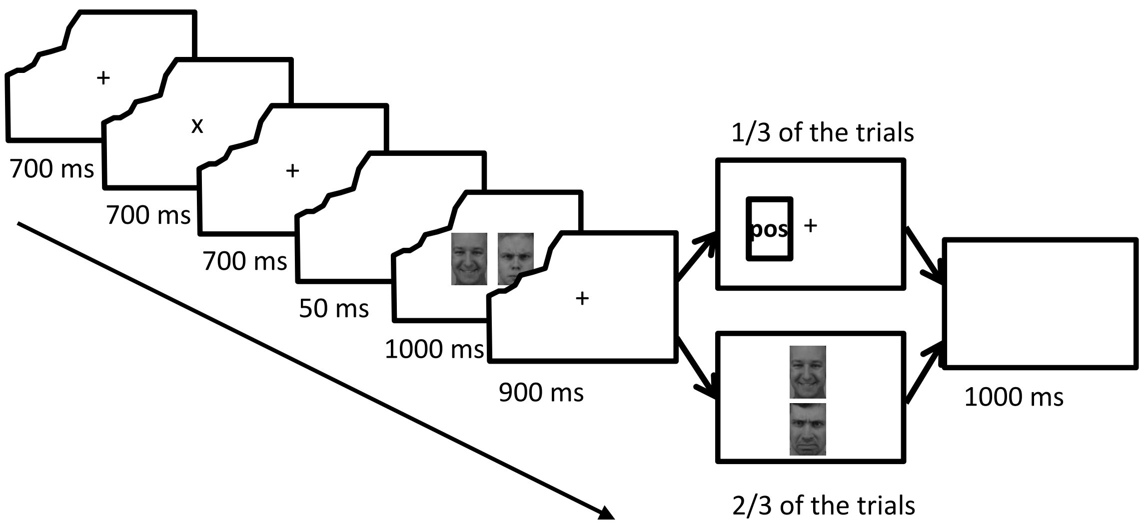


Figure 15. *Trial sequence of the change detection task in Experiment 3b. (Depicted is a change trial with respect to the identity and a trial that requires a no change response regarding the emotion).*

Due to the simplified trial sequence compared to Experiment 3a, the practice phase was also shortened. First, there were 8 trials in which the identity change detection task was practiced, followed by 8 practice trials for the valence change detection task. After that, there was a final practice block with 24 practice trials in which participants did not know during study and maintenance whether the identity or the valence of a face would be tested (in these practice trials two-third of the trials, which is 16 trials, tested the memory for identities and in one-third, which is 8 trials, memory for the face valence was tested). After this last practice block, the main phase started.

Again a block-feedback showing the percentage of correct responses (for both valence change detection and identity change detection) was presented every 48 trials (and after each block of the practice phase). In this experiment, in which we tried to obtain higher accuracies, participants were encouraged to try to remember the faces better if performance was below 85 percent in the identity change detection trials. If performance in the valence change detection trials was below 60 percent, they were encouraged to engage more in this task.

B2.2 Results and Discussion

The results of Experiment 3a-c are reported in a common section. For Experiment 3b, there was significantly better performance in incongruent trials compared to congruent trials measured with *d'*. The analysis of reaction times revealed faster responses in congruent compared to incongruent trials, which did not interact with the factor change type. Thereby, Experiment 3b has a clearly diverging pattern of results in the measures *d'* and response times. Although participants were able to discriminate better between changes and no changes in incongruent trials, (correct) responses were faster in congruent compared to incongruent trials. There are two central differences that might have caused differences in results between Experiment 2a-d, in which an beneficial effect of evaluative congruency was observed, and the current studies, Experiments 3a and 3b. Besides the locational changes, another central difference is the presence of trials that introduced a strong focus on the valence of the faces. This last point was addressed in Experiment 3c.

B3 Experiment 3c

The influence of congruency on reaction time and accuracy measured with *d'* pointed in different directions in the previous experiment. We reasoned that the trials with a change detection task focusing on valence might have introduced a further process and a possible confound into the design: To obtain good performance in the change detection task focusing on valence, in congruent trials, participants could only store *one* valence, which is the same for both faces. They could ask themselves "Are both faces positive or negative?" Only storing the information "negative" (or "positive") in one trial can lead to a perfect performance in the valence-change detection task in congruent trials. In incongruent trials, in contrast, participants have to remember *which emotion* was presented *at which location*. E.g. they could remember "the *right* face was *positive* (and the left one was negative)". Thus, at least two pieces of information have to be stored in incongruent trials. Therefore, this could have artificially led to a higher load in incongruent trials. Probably, this could have led to the slower responses in incongruent trials compared to congruent trials observed in the previous experiments (Experiment 3a-b). Therefore, the strong focus on valence was deliberately given up in this experiment.

If these speculations turn out to be true, the congruency effect in reaction times should vanish in the current experiment. However, the question remains, whether the focus on valence can also explain the reversed effect of congruency on accuracy measured with *d'* that is in contrast to the finding in Experiment 2a-d. It could be speculated that a specific *focus* on similarities could lead to higher confusability in working memory paradigms. E.g. while Davelaar and colleagues (2006), who used only weak associates, observed beneficial effects of semantic similarity, some other studies using more obviously and strongly associated material observed no effect or mixed results (Baddeley & Levy, 1971; Cowles et al., 2010, Experiment 2; van der Lely & Howard, 1993) or even a reduced memory performance for related stimuli (Baddeley, 1966; Dale & Gregory, 1966). However, this assumption remains rather speculative.

B3.1 Methods

A total of 70 students (46 females, age range: 18-35, *Md* = 23) with normal or corrected to normal vision participated in Experiment 3c and were paid 4 € for participation. Again, it can be assumed that with $N = 70$ participants and $\alpha = .05$ (two-tailed) a medium effect of $d_Z = .35$ would be detected with $1\text{-}\beta = .90$.

Design and Procedure. Again a change detection task with locational changes from study to test was used. The essential manipulation was the manipulation of evaluative congruency.

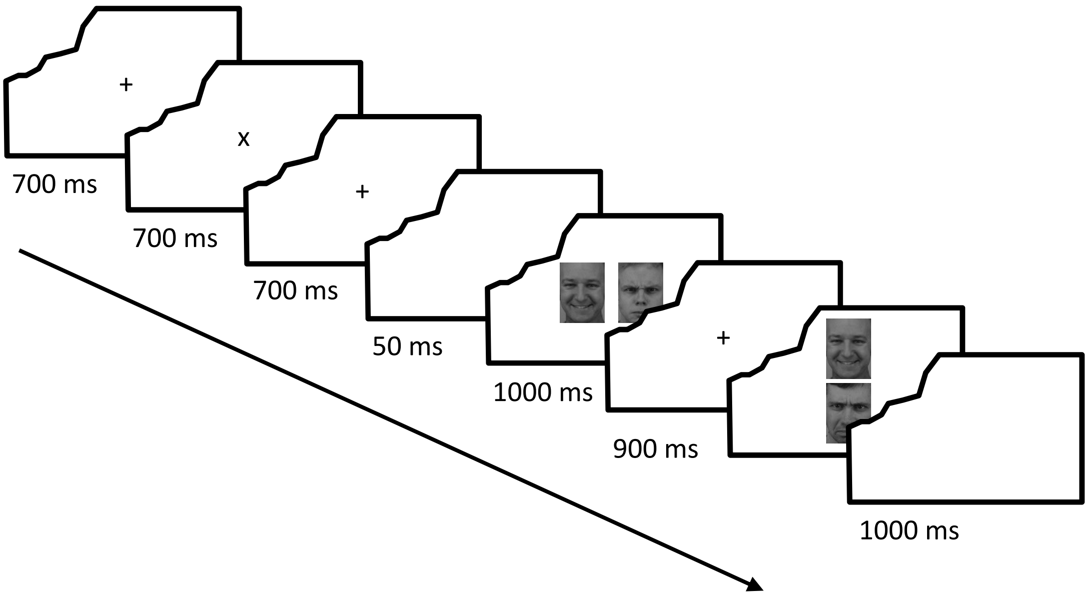


Figure 16. *Trial sequence of the change detection task in Experiment 3c. (Depicted is a change trial).*

The procedure was the same as in Experiment 3b with the exception that the trials with the focus on valence were dropped (see Figure 16). Therefore, this experiment was shortened in its duration compared to Experiment 3b but it still contained 128 relevant trials with a change detection task in which the memory for the identities of the faces was tested. Correspondingly, compared to Experiment 3b the practice phase was shortened. There was only a single practice phase in which participants were familiarized with the identity change detection task, consisting of 8 trials. The feedback and the blocked feedback for the identity change detection trials remained the same as in Experiment 3b.

B3.2 Results and Discussion

The results of Experiment 3a-c are reported in a common section. In Experiment 3c, using *d'*, a higher performance in incongruent trials compared to congruent trials was observed. In the common results section, a main effect indicating faster correct responses in congruent compared to incongruent trials is reported that was, however, significantly moderated by the factor change type. Further investigation of this interaction revealed that in change trials, responses were faster in congruent compared to incongruent trials, while for no change trials, there was no significant effect of congruency. Thus, in reaction times of correct responses, a reversed pattern compared to the analysis with *d'* was again observed. This pattern was however restricted to change trials.

Primarily, it can be concluded that rather similar results were obtained irrespective of whether a focus on valence was introduced or not. Without the focus on valance and without valence being in any way task-relevant, the potential problem that an additional load is introduced in incongruent trials is removed. Nevertheless, responses occurred faster in congruent compared to incongruent trials, at least in change trials. Therefore, it seems that using this procedure with locational changes of the items from study to test, hints of a speed-accuracy tradeoff can be obtained.

To finally answer the question whether the effect in the accuracy data of Experiment 3a-c or the effect in reaction times is a more reliable measure of working memory performance in this task a drift-diffusion model analysis for all three experiments (Experiment 3a-c) was performed (see section 4.1.4).