RUNNING HEAD: CROSS-MODAL ENHANCEMENT OF SEARCH EFFICIENCY

Semantically congruent auditory primes enhance visual search efficiency: Direct evidence by varying set size

Timea Folyi, Michaela Rohr, Lisa Peuckmann, &
Dirk Wentura

Saarland University, Saarbrücken, Germany

-------- published in Journal of Experimental Psychology: HPP, 2025 -------

The final article is available via its DOI: https://10.1037/xhp0001252

Abstract

Task-irrelevant sounds that are semantically congruent with the target can facilitate performance in visual search tasks, resulting in faster search times. In three experiments, we tested the underlying processes of this effect. Participants were presented with auditory primes that were semantically congruent, neutral, or incongruent to the visual search target, and importantly, we varied the set size of the search displays. According to seminal accounts of semantic priming, priming effects can be explained by processes not related to search (i.e., facilitation of target encoding; McNamara, 2013), which would predict a priming effect that is independent of set size. Alternatively, we tested if auditory priming can serve as a source of guidance for visual attention toward the primed target (i.e., in terms of altering attention-directing priorities; Wolfe, 2021), as indexed by higher search efficiency with congruent priming. Experiment 1 found that auditory color word primes resulted in faster responses and, importantly, flatter search slopes for congruent compared to incongruent color targets, indicating more efficient search. As with many naturalistic search behaviors, we used multiple-target search. Experiment 2 replicated the findings of Experiment 1 with a reduced target set. Experiment 3 extended these findings to complex audio-visual objects. Our results provide direct evidence that cross-modal priming can guide visual selective attention, as reflected by enhanced visual search efficiency.

Keywords: priming, visual search, cross-modal, attention

Public Significance Statement

Sounds can be beneficial when searching for a visual object if they are meaningfully related to the object we are looking for. Our study investigated the basic cognitive mechanisms of this effect. We tested if target-congruent sounds can guide visual attention, or if they provide benefits related to target identification, but unrelated to search. Searching for the target became more efficient when the task-irrelevant sound denoted the target, rather than a potential but absent target. This result supports that congruent sounds guide visual attention. The findings contribute to our understanding of how auditory warning signals influence visual search behavior.

Author Note

Timea Folyi (e-mail: t.folyi@mx.uni-saarland.de), Michaela Rohr (e-mail: m.rohr@mx.uni-saarland.de), Lisa Peuckmann (e-mail: lisa.peuckmann@uni-saarland.de), and Dirk Wentura (e-mail: wentura@mx.uni-saarland.de). The affiliation of all authors is Department of Psychology, Saarland University, Saarbrücken, Germany. We would like to thank Ullrich Ecker for his helpful comments on an earlier version of this article. Correspondence concerning this article should be addressed to Timea Folyi, Department of Psychology, Saarland University, Campus A2 4, D-66123 Saarbrücken, Germany, e-mail: t.folyi@mx.uni-saarland.de.

The data and materials for all experiments are available at https://osf.io/qse53/?view_only=e98681d7e2f84d74a18e4e66ddaf64bc. The preregistration of Experiment 1 is available at https://aspredicted.org/blind.php?x=hc26e5; the preregistration of Experiment 2 at https://aspredicted.org/FLD_867; and the preregistration of Experiment 3 at https://aspredicted.org/blind.php?x=vh99jg.

Semantically congruent auditory primes enhance visual search efficiency: Direct evidence by varying set size

In natural environments, the human cognitive system permanently receives and integrates information from different senses. Since visual and auditory information often occur together in our natural surroundings, it is particularly likely that visual and auditory information are processed and integrated cross-modally. While initially somewhat neglected by cognitive research, cross-modal information processing has received increasing attention over recent decades, whereby its distinctive features compared to unimodal information processing are becoming more and more apparent (e.g., Calvert et al., 2004; Driver & Spence, 1998; Spence et al., 2009). For example, converging multisensory information is readily integrated at multiple stages of information processing, leading to marked performance benefits that cannot be explained by a simple accumulation of the effects of each modality (e.g., De Gelder & Bertelson, 2003; Giard & Peronnet, 1999; Talsma et al., 2010). In addition, the specific characteristics of each modality could lead to benefits by increasing the strength and reliability of the converging information. For example, due to the unique characteristics of auditory perception, such as its near omnidirectional character, sounds might be highly effective in conveying critical information when the visual environment is complex and cluttered (e.g., Fernandez-Duque & Posner, 1997).

Since semantic information is represented primarily in the visual and auditory domains, one obvious research question concerns the potential of auditory stimuli to facilitate the processing of semantically congruent visual stimuli. There are now many studies showing facilitated processing of visual stimuli if their presentation is accompanied by a semantically congruent auditory stimulus (e.g., Chen & Spence, 2011, 2013, 2018a, 2018b; Holcomb & Anderson, 1993; Tabossi, 1996). These studies have used variants of the well-known semantic priming paradigm, which was established in the visual domain (for reviews, see McNamara, 2005, 2013). Typical semantic priming studies present the target stimulus at a central, predetermined and therefore attended location. Priming effects (i.e., semantically related primes lead to faster responses to the target and/or less errors compared to unrelated primes) are then explained by facilitation of target encoding. To illustrate, the most prominent explanation (Collins & Loftus, 1975) assumes that the prime activates its mental representation, which in turn pre-activates the mental representation of the congruent target through a process of “spreading activation”. Hence, the attended target stimulus is more easily encoded (for detailed discussion, see McNamara, 2005, 2013).

However, in contrast to the well-established paradigm with predefined and thus attended target location (see e.g., Chen & Spence, 2011, 2013, 2018a, 2018b; for cross-modal studies), several cross-modal priming studies are characterized by uncertainty of the target location within a cluttered visual context. Thus, the target is embedded in a set of distractors (e.g., Iordanescu et al., 2008, 2010, 2011; Mahr & Wentura, 2014, 2018). Although this setting may seem odd in the context of the predominantly unimodal research on semantic interactions, it is consistent with natural audio-visual information processing. When searching for (often multiple or broadly defined) targets in complex visual scenes, the salient, near omnidirectional auditory input could effectively inform the visual system about meaningfully related, potentially relevant items. Obviously, the involved processes might also differ between these settings.

When the target is embedded in a complex visual scene, a visual search process takes place before elaborate target encoding can proceed. Drawing from the established guided search model of selective visual attention (Wolfe, 2021), various sources of guidance inform the search process, including not only bottom-up physical salience but also current selection goals, prior information (e.g., selection history), values, and learned regularities of the visual scene. Visual working memory or long-term memory is thought to mediate these latter influences (Wolfe, 2021). The weighted average of the various contributions is integrated into a representation of the attention-directing priorities of the visual scene, the priority map. Based on this model, attention is guided, covertly or overtly, to the item or location with the highest activity on the priority map. The representation of attentional priorities is dynamic, thus, current attention-guiding influences update the priority map constantly (e.g., Wolfe, 1994, 2021).

One prominent source of attentional guidance is the current selection goal. Specifically, this involves attentional guidance based on the mental representation of the target-defining features, often referred to as the attentional template (e.g., Duncan & Humphreys, 1989) or guiding template in the terminology of guided search (Wolfe, 2021). It is assumed that this source of guidance is exerted dominantly via working memory (e.g., Grubert & Eimer, 2018; Kerzel & Witzel, 2019; Olivers et al., 2011), and it has a direct influence on the priority map (Wolfe, 2021). According to the guided search model, further sources of guidance (e.g., repetition priming; Kristjansson & Johannesson, 2014) can implicitly influence the prioritization of target features in the guiding template (Wolfe, 2021). In a similar vein, auditory semantic priming might also be considered as a source of guidance for visual attention toward the primed target, tentatively, by affecting the guiding template of the search process. Consequently, the target-congruent information conveyed by sounds may alter the distribution of activation in the priority map in favor of the primed target, enhancing its visual salience.

But how could the semantic information provided by sounds affect the priority map of visual attention?

The "good-enough" principle of attentional guidance (Yu et al., 2023; see also Wolfe, 2021) stands for the notion that visual attentional guidance is based on a "coarse" representational quality of the target-defining features, that is "good enough" for rapid localization of potential targets. These representations serve as the basis for sensory gain modulation, thereby influencing activation of the priority map (Wolfe, 2021). Despite extensive research, the features that can be represented in guiding templates are still not sufficiently specified, especially in the context of realistic search displays. Color, shape, and orientation properties, and low-spatial-frequency information are candidates regarding guidance to complex objects (Alexander et al., 2019; Wolfe, 2021; Zhang & Li, 2020; see also Van Moorselaar & Theeuwes, 2023, for proposing object-based attentional prioritization). Importantly, there is evidence that guiding templates can be set up based on complex semantic information: For example, when searching for semantic categories, they can represent the typical features of these categories (e.g., Maxfield et al., 2014; Yu et al., 2016). Moreover, guiding templates can be influenced by the semantic context, such as complex scene information. For example, visual search for the semantic category "bottle" will be guided by the object features that characterize a beer bottle when primed with an image of a bar scene before the search display, and by those of a baby bottle when primed with an image of a baby's room (Robbins & Hout, 2020). On a similar vein, auditory semantic priming may act as a source of guidance for visual attention by facilitating sensory gain for a primed feature, when it is sufficient to define a target, or, in the case of priming of object representations, for a coarse representation of the primed target object's features, similar to search based on word cues or categories.

After a rapid localization of a potential target candidate, a more elaborate processing follows that enables item identification, which informs the decision if the selected item matches a target representation (Wolfe, 2021; Yu et al., 2023). Therefore, after selection has occurred, there is room for a possible facilitation by semantic priming in terms of spreading activation processes (e.g., McNamara, 2013) to support identification of the item. The decision whether the identified item is a target is based on a more veridical target representation than the information used for guidance (i.e., "target template" of the activated long-term memory in the terminology of guided search; Wolfe, 2021).

In the present study, we target possible underlying processes of cross-modal semantic priming in visual search tasks. Specifically, based on seminal accounts of semantic priming, a performance advantage by congruent (as opposed to incongruent) priming can be explained by better accessibility of semantic information during target processing (e.g., McNamara, 2013), implying that primes exert an influence on processes following selection (i.e., target identification, target-match decision). In contrast, we propose that a facilitation that is directly related to attentional guidance might also be involved: Cross-modal semantic priming might serve as a source of guidance for visual attentional selection in favor of the primed target (see, Wolfe, 2021). That is, an auditory prime might increase the relative salience of the congruent target in the cluttered visual context, so that attention can be more efficiently directed to it.

Importantly, these two accounts can be tested in our cross-modal visual search experiments via variation of the number of elements on the search display, that is, the set size. In basic visual search tasks, the set size is systematically related to the reaction time: The time required to find the target increases (typically) linearly with increasing set size (e.g. Wolfe, 2014). A robust, well-replicated marker of search efficiency is the gradient of the search slope, that is, the increase in reaction time per additional distractor item (e.g., Wolfe, 1998, 2021; Wolfe & Horowitz, 2017)—the flatter the slope, the more efficient the search. Thus, if cross-modal priming involves guiding of visual attention, one would expect more efficient search, that is, flatter slopes with congruent compared to incongruent priming. Remarkably, as indicated by our review of cross-modal priming studies involving visual search (i.e., uncertainty of the target location) in the following section, studies with set size variation are lacking so far. Thus, the aim of the present study was to test directly, by varying the set size, if cross-modal semantic priming leads to more efficient search in case of prime-target congruence.

Cross-modal semantic priming in a search context

Without addressing the involved processes directly, there are several studies providing hints that auditory semantic priming of visual search might not only facilitate processes not related to search, such as encoding facilitation of the target object, but also increase visual search efficiency. By using naturalistic sounds as primes, studies by Iordanescu et al. (Iordanescu et al., 2008, 2010, 2011; see also Zweig et al., 2015; Salverda & Altmann, 2011; and Huettig & Altmann, 2005) investigated cross-modal semantic priming in a visual search context.[1] For example, Iordanescu et al. (2008)

[1] Of note, cross-modal exogenous cuing is also a task that includes uncertainty regarding the spatial location of the relevant stimuli, but presenting a very limited visual display. Using this task, the study of Mastroberardino et al. (2015) found limited evidence for cross-modal semantic congruency effect. There was a significant effect of semantic congruency only when using a difficult perceptual discrimination task and when presenting target displays with two items. However, there was no effect of semantic congruency when only one target item was presented.

presented four pictures of objects in each trial. The target object (e.g., a cat) that had to be located in the current trial was announced as a spoken denotation at the beginning of each trial (i.e., about 1 second before the search display). Synchronously to the search display, a complex, naturalistic sound was presented that was either a characteristic sound of the target object (e.g., a “meow”) or not. It was indeed found that target-related sounds sped up responses. Iordanescu et al. (2010) conceptually replicated this effect with saccadic search times as the dependent variable. Kvasova et al. (2019) extended the findings for a complex search task that resembles dynamic, naturalistic scenes: Target-congruent auditory primes (again, naturalistic sounds) reduced response times compared to control conditions. Taken together, Iordanescu et al. (2008, 2010, 2011) and Kvasova et al. (2019) discuss their results in the context of attentional facilitation by congruent priming: Specifically, cross-modal interactions in object processing might increase the visual salience of the target object. Without testing the influence of primes on the efficiency of search processes, however, the processes involved are unclear: They might include improved search efficiency in a cluttered visual context, or they might be limited to post-search processes, in particular, facilitating the encoding of an attended item and/or reducing the time taken to accept or reject an item as target.

Since semantic priming by naturalistic sounds may involve somewhat different mechanisms than priming by spoken words (e.g., Chen & Spence, 2018b; however, see also Iordanescu et al., 2011, for comparable auditory–visual congruency effects for these types of stimuli), studies with spoken word primes are more directly relevant to the present context. Nevertheless, even in this line of research no variation in set size has been used. For example, Mahr and Wentura (2014; Exp. 3) instructed participants to make target absent/present decision for briefly presented visual search displays of colored circles. Participants had to indicate whether one of the four possible target colors (blue, green, red, yellow) was present. The target circle (if present) was presented at a random location in a ring-like arrangement along with seven other circles, colored in heterogeneous non-target colors. In the majority of trials, visual displays were preceded by an auditory prime, thus, a spoken word naming one of the target colors, or a neutral word. Importantly, the prime was not predictive of the actual target color (i.e., primes and targets were varied orthogonally). Discrimination sensitivity was significantly increased in the congruent condition compared to incongruent, neutral, or no-prime conditions. As mentioned above, there was no variation of set size to test the process of this performance facilitation. Furthermore, Mahr and Wentura (2018; Exp. 2 and 3) extended these findings to an applied context: Participants took part in a simulated driving task with target symbols (e.g., a traffic light symbol) appearing on overhead gantry signs. As in Mahr and Wentura (2014), there were four different possible targets, but on each trial, only one target was presented at one of four possible locations, together with three distractors. Before target onset, spoken primes named one of the four possible targets (e.g., “traffic light”). These spoken denotations were uncorrelated with the identity of the target that followed them. Participants had to categorize the color of the target. Auditory primes that named the target facilitated response times compared to incongruent and neutral primes. Interestingly, congruently (vs. incongruently) primed targets were not only associated with lower mean RTs, but also lower mean SDs of raw RTs. While the results (similar to those of Mahr & Wentura, 2014) are discussed in the context of enhanced target processing by congruent priming, the pattern of results also fits the assumption that congruent priming might increase search efficiency, as inefficient search is characterized by large variability in the number of search steps required to find the target. Again, there was no variation in set size to test this hypothesis.

Furthermore, a potentially relevant difference between the experiments of Mahr and Wentura (2014, 2018), on the one hand, and the line of research of Iordanescu et al. (Iordanescu et al., 2008, 2010, 2011) and Kvasova et al. (2019), on the other hand, concerns the search process involved in the task. In particular, the latter line of studies announced the target of the current trial (i.e., as spoken or written denotation) selected from a large set of potential targets shortly before the auditory prime was presented (Iordanescu et al., 2008, 2010, 2011; Kvasova et al., 2019; see also Knoeferle et al., 2016). Since the specific target of the current trial was clearly defined, the involved process can be interpreted as single-target search. However, this single-target approach had a potentially problematic aspect: As a new attentional template had to be set up for each search trial, the auditory primes may have already influenced this process, making it difficult to accurately gauge the underlying processes. Thus, in line with the experiments by Mahr and Wentura (2014; 2018), we decided to adopt a task involving multiple-target search. As with many naturalistic search behaviors, this setting involves a small set of potential targets and requires participants to maintain the possible target items across multiple search trials.

Altough the mentioned studies of basic attentional research did not vary set size within an experiment, there is a related study that gives further hints for facilitation of search efficiency by cross-modal priming: Knoeferle et al. (2016) used the paradigm developed by Iordanescu et al. (2008) in an applied setting with a set size variation between participants. In their Experiment 4B, participants' task was to search for a specific product in one versus three rows of six products on an *Amazon*-style product selection page. The authors found significant facilitation of RTs for the product primed by a characteristic sound only in the condition with three rows—a result that can be interpreted, with some caution, as increased search efficiency. However, due to the applied research question and the goal of the study to test the influence of "visual load" on priming of product search (i.e., one versus three rows of products were introduced as low versus high visual load conditions, respectively; see e.g., Lavie, 2010), it was not the intention of the authors to mimic a typical visual search experiment. Hence, the experiment was a single-trial study to resemble a realistic online shopping context, making it difficult to relate the results to typical visual search research.

In sum, several studies have integrated visual search with a cross-modal semantic priming design, and they suggest that priming effects may arise, at least partly, from increased search efficiency. However, none of the studies in basic visual search research has included a variation of set size to test this assumption directly. Apart from studies using variants of semantic priming in a visual search task, some other paradigms have provided further hints that auditory primes may affect visual search processes. These studies are discussed below.

Further insights from related paradigms

Cross-modal response competition paradigms and perceptual load

A number of studies have used cross-modal response-competition paradigms[2] with visual targets and auditory flankers (e.g., Mahr & Wentura, 2014, Exp. 1 and 2; Mahr & Wentura, 2018, Exp. 1; Tellinghuisen & Nowak, 2003; Tellinghuisen et al., 2016). These studies found large congruency effects; that is, responses were faster if the

[2] We use "response-competition paradigm" here as an umbrella term for flanker (Eriksen & Eriksen, 1974) and Stroop tasks (Stroop, 1935); these paradigms share the characteristic that distractor information (i.e., the flanker stimulus or the irrelevant color word) is either compatible or incompatible with the required target response.

auditory flanker matched the visual target compared to incongruent pairings. Of course, flanker effects are usually interpreted in terms of response facilitation/interference due to a response-relevant feature overlap between flanker and target, therefore, it is not straightforward to conclude if the experiments involved other than solely response-based effects. Of most relevance, in some of these studies (Tellinghuisen & Nowak, 2003; Tellinghuisen, et al., 2016; Mahr & Wentura, 2014), the response-competition task was combined with a perceptual load manipulation (Lavie, 1995, 2005). For example, in Mahr and Wentura (2014), participants were presented with the colored target circle in a ring-like arrangement including seven other circles colored in either different non-target colors (i.e., high perceptual load) or gray (i.e., low perceptual load). This load manipulation might be compared to a visual search manipulation with set size = 1 (low load, because of the uniformly grey circles) versus set size = 8 (high load). Thus, with some caution, the increase of the response-compatibility effect in high load condition might be interpreted as hints for increased search efficiency by congruent priming.

Cross-modal correspondence and the pip-and-pop effect

On a related note, Klapetek et al. (2012) investigated the influence of a natural cross-modal correspondence (see Spence, 2011, for a review) between primes and targets on visual search processes. It is an established effect in cross-modal attention that sounds can facilitate visual search due to temporal coincidence: When a visual target has unique dynamics compared to the distractors and a sound is synchronized to this change, visual search efficiency is enhanced (i.e., “pip-and-pop effect”, Van der Burg et al., 2008). Klapetek et al. (2012) tested this effect with varying brightness (light vs. dark) as the dynamics, and most importantly, they tested whether this effect was enhanced by a cross-modal correspondence with the pitch of an irrelevant tone. Hence, the authors presented congruent (i.e., the low-pitch/high-pitch beep synchronized with the darker/brighter target) and incongruent pairings (i.e., the reversed assignment). There were no congruency effects with trial-by-trial variation of congruency. When congruency was varied block-wise, mean response times were lower for congruent blocks; however, search efficiency was not affected by congruency.

Real-time influence of spoken instructions on visual search

A further line of research provided evidence that task-relevant verbal information can have an immediate, real-time influence on visual search processes (e.g., Reali et al., 2006; Spivey et al., 2001). In the study of Spivey et al. (2001), simple bar stimuli served as targets with certain color by orientation conjunctions, among distractors with different color by orientation conjunctions. Participants received spoken instructions simultaneously with the visual search display; they denoted the first, then the second feature that identified the target (e.g., “Is there a red vertical?”). Incremental processing of spoken instructions resulted in more efficient search than might have been expected given the difficulty of the conjunction search. This suggests that participants were able to rapidly use the consecutively presented feature adjectives, resulting in two consecutive but more efficient search processes. While this research demonstrated that target-relevant spoken and visual information can be rapidly integrated during visual search, evidently, the auditory information is used strategically here since it specifies the target of the current trial. Therefore, this instance of search facilitation can be clearly delineated from task-irrelevant influences on search.

In sum, studies using other paradigms than semantic priming can provide some hints with regard to our hypothesis of increased search efficiency by congruent auditory primes: The studies using response-competition paradigms (e.g., Mahr & Wentura, 2014; Tellinghuisen & Nowak, 2003) suggest an affirmative answer regarding the involvement of search facilitation. However, the interpretation of a perceptual-load manipulation as a set-size variation is only post-hoc, and the response-competition design is not optimal for studying visual search processes due to an assumed predominance of response-related processes in this task. Furthermore, the exploration of whether cross-modal correspondences increase search efficiency in the pip-&-pop paradigm (Klapetek et al., 2012) yielded a negative answer. Again, these results should be interpreted with caution in the present context, as the task arguably involves different processes. Finally, studies of linguistically mediated search (e.g., Spivey et al., 2001) suggest a positive answer, demonstrating that target-relevant spoken and visual information can be rapidly integrated during visual search, at least when linguistic information is used strategically.

The present study

The present study aimed to directly test the underlying processes of cross-modal semantic priming in a visual search task. To this end, we implemented visual search experiments accompanied by auditory priming, in which the set size of the search displays was varied. Therefore, we presented visual search displays with varying number of items and spoken word primes that were semantically congruent, incongruent or neutral to the search target. The primes contained no spatial information (i.e., they were presented centrally on each trial) and they were non-predictive in the sense that primes and targets were combined orthogonally, so that they did not provide any predictive information about the identity of the target on a given trial. Moreover, the primes were not response-related, as the participants' task was to categorize a visual feature of the target that is varied orthogonally to the priming. We tested the involvement of search facilitation, first, by using simple search displays characteristic of basic visual search research (Experiments 1 and 2). Second, we tested if the assumed involvement of search facilitation extends to more complex audio-visual objects that have relevance to search behavior in real life (Experiment 3).

Specifically, in Experiments 1 and 2, we used a perceptual feature to define the targets, thus, participants had to search for circles with colors from a predefined target color set. Non-targets were circles with heterogonous colors that were not part of the target set (*Figure A1* of the *Appendix;* see also Mahr & Wentura, 2014). Participants were presented with search displays of set size 2, 8, or 16. In all trials, there was one circle with a target color. Auditory primes (i.e., spoken color words) accompanied the visual presentation. The task was to classify a wedge-shaped gap in the target circle as facing up or down (see *Figure 1*), thus, response-priming processes can be ruled out. Similar to many naturalistic search behaviors, we used a task with multiple-target search (see also Mahr & Wentura, 2014, 2018). While Experiment 1 featured four potential targets, Experiment 2 featured a reduced set of two potential targets, thus, reduced cognitive load. Furthermore, Experiment 3 aimed to test our hypothesis with perceptually and semantically more complex stimuli that have more relevance to real-life search processes. Specifically, we presented automotive symbols as targets (e.g., traffic light) and distractors (i.e., heterogeneous automotive symbols; see *Figure 4*), while spoken denotations of potential targets served as primes (see also Mahr & Wentura, 2018). That is, targets were not defined by a simple perceptual feature but by their semantic meaning. Again, we varied the size of the search displays between 2, 8 and 16 items. To rule out response-priming processes, the task of the participants was to categorize the orientation of the target symbol as left or right. Hence, we were able to test whether the hypothesis about facilitation of visual search by cross-modal priming generalizes to complex cross-modal object representations.

Overall, we expected priming effect, thus, faster target responses with congruent versus incongruent primes. Furthermore, we expected slower response times with increasing set size: Although targets in Experiments 1 and 2 were defined by a single feature, we expected a linear relationship between search times and set size given the highly heterogonous non-target colors (e.g., Duncan, 1989; Nagy & Sanchez, 1990). Most importantly, we aimed to test if search slopes are flatter for congruent compared to incongruent trials, thus, if congruent primes are associated with more efficient search. The prevailing account of semantic priming effects, that is, encoding facilitation of the attended target (e.g., McNamara, 2005, 2013) predicts a priming effect regardless of the set size. However, if auditory priming serves as a source of guidance for visual attention (i.e., in terms of biasing the priority map of the visual scene in favor of the congruent target; see Wolfe, 2021), flatter search slopes are expected with congruent compared to incongruent priming.

Following Mahr and Wentura (2014, 2018), we used time-compressed auditory stimuli, which allow implementation of presentation times and prime-target stimulus onset asynchronies (SOAs) that are typical of visual priming studies (i.e., in the present study: SOAs of 90 and 100 ms), and therefore increase the comparability of these lines of research. While time-compressed auditory stimuli allow for nearly immediate effects on visual search, previous results have shown that compression itself or the degree of compression has no influence on semantic congruency effects as long as the primes are clearly understandable (Mahr & Wentura, 2014, 2018). As all possible prime-target combinations were presented equally often, primes were entirely non-predictive regarding the identity of the actual target. Note that demonstrating cross-modal priming effects with brief prime durations and prime-target SOAs, and with non-predictive primes would support "automaticity" of the involved processes in the sense of their speed and involuntariness (e.g., Moors, 2016). We thus regard these features as strengths of the present design.

Experiment 1

Method

Transparency and Openness

The experiment was preregistered at https://aspredicted.org/blind.php?x=hc26e5. In the present study, there were some minor deviations from the preregistrations. We list and explain them in *Table A2* of the *Appendix.* For all experiments, we report details on the determination of sample size as well as all data exclusions, manipulations, and all measures in the study. The data and materials for all experiments are available at https://osf.io/qse53/?view_only=e98681d7e2f84d74a18e4e66ddaf64bc.

Participants

We based power calculations on the effect sizes found by Mahr and Wentura (2014). In Experiment 1, this effect size was $d_Z = .67$; in Experiment 2, it was $d_Z = .61$. Since we reduced the number of trials (from 300 to 240) and changed the task, we reduced the expected effect size to $d_Z = .50$ (i.e., medium effect as defined by Cohen, 1988). To detect an effect of this size with power $1 - \beta = .95$ ($\alpha = .05$, one-tailed[3]), a sample size of $N = 45$ is needed.

Forty-five students (29 women, 16 men) from Saarland University took part in the experiment in exchange for 4 Euro. The median age was 24 years (range 18 to 33). All participants were native speakers of German and had self-reported normal hearing and normal or corrected-to-normal vision. The data of five further participants were discarded because error rates exceeded the pre-registered criterion (three interquartile ranges above the third quartile with respect to the sample distribution; *far out values* as defined by Tukey, 1977).

Design and Statistical Analyses

The experiment used a 3 (prime-target relation: congruent, incongruent, neutral) × 3 (set size: 2, 8, 16 stimuli) within-participants design—although technically, the congruency factor was implemented by fully crossing auditory prime (red, green, blue, yellow, neutral) and visual target (red, green, blue, yellow) factors. Thus, prime and target features were uncorrelated, resulting in auditory priming without contingency (i.e., there was no benefit in expecting the primed color).

In line with our preregistration, we tested our hypotheses by a 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) multivariate analysis of variance (MANOVA) for repeated measures (e.g., Dien & Santuzzi, 2005; O'Brien & Kaiser, 1985). Additionally to the tests that include the factor of prime-target relation (i.e., with $df = 2$), we report the results for the contrast between congruent and incongruent conditions, thus, the contrast that is most central regarding our hypotheses

[3] One-tailed interpretation because of the expected positive congruency effect on RTs, thus, faster responses in trials with congruent priming compared to trials with incongruent priming.

(see also Mahr & Wentura, 2014, 2018). The set size factor was transformed into linear and quadratic trends: The linear trend is of main interest, since we expect a linear relationship between set size and search RTs.[4]

To test the influence of priming on search slopes more directly, as preregistered, we report linear mixed model (LMM) analyses, which account for the hierarchical structure of the data, in which trials (level 1) are nested within participants (level 2). Specifically, we analyzed RTs as a function of priming condition (coded as +1 congruent, -1 incongruent), set size (i.e., 2 / 8 / 16), and their interaction. The weight for display size plus the weight for the interaction is the search slope for congruent trials; the weight for display size minus the weight for the interaction is the search slope for incongruent trials. The test for the interaction is the test for differences in search slopes.

Material

Each target display contained two, eight, or sixteen visual stimuli presented in random locations of a 4 × 4 matrix (18.5 × 18.5 cm, approx. 18.5 × 18.5 ° visual angle) on a black background (see *Figure 1*). Each circle spanned 3.6° (diameter 73 pixels = 3.6 cm). One of these visual stimuli was the target and therefore appeared in one of the four target colors (i.e., blue, green, yellow, red; see *Figure A1* of the *Appendix*). The filler items were presented in different non-target colors, randomly chosen from 16 filler colors that were clearly distinguishable from the target colors (see *Figure A1*). All circles had a gap in one of six positions, three in the upper part of the stimulus (-45°, 0°,

[4] The multivariate approach to the repeated measures analysis transforms a tripartite factor into a vector of two orthogonal contrast variables (see, e.g., Dien & Santuzzi, 2005). For the prime-target relationship, the first contrast was chosen a priori as the contrast between congruent and incongruent trials (i.e., the most relevant comparison). Consequently, the second contrast compared the average across congruent and incongruent trials with the neutral trials. Obviously, this contrast is of little interest in the present study. For the set size factor, we used polynomial contrasts. Of course, in the case of three levels, the linear trend is simply the comparison of the first level with the last level, whereas the quadratic trend refers to the comparison of the middle level with the average of the first and second level.

45° with reference to the 12 o'clock position) and three in the lower part (-45°, 0°, 45° with reference to the 6 o'clock position).

We used the auditory word stimuli of Mahr and Wentura (2014), that is, time-compressed sound files (with 120 ms duration, compression to 30% of their original length; thus, the primes of Exp. 3, and the medium compression rate of Exp.1-2 in Mahr & Wentura, 2014) of the German one-syllable words for red ("rot", [ro:t]), green ("grün", [gry:n]), blue ("blau", [blaʊ]), and yellow ("gelb", [g ɛlp]). Neutral primes were also chosen from a set of four possible stimuli (see Mahr & Wentura, 2014, Exp. 2): Four one-syllable non-words served as neutral prime set ("liez", [li:ts], "tän", [tɛ:n], "nux" [nʊks], and "töff" [toef]).[5] The sounds were presented over closed-ear headphones (model AKG K511) and ranged in sound pressure level from 68 to 72 dB SPL. The 30%-compression files were clearly understandable (i.e., identification accuracy was about 100%, as reported by Mahr & Wentura, 2014).

Procedure

Participants were individually seated in front of a 17-inch monitor (100 Hz refresh rate, resolution 640 × 480 pixels) controlled by a personal computer in a sound-proof testing booth with dimmed light. Viewing distance was about 60 cm. The experiment was controlled by E-prime software (E-prime 2.0). On each trial, participants had to search for the target color and categorize the target-gap location as up or down by pressing a corresponding key (keys F and J on a standard QWERTZ keyboard;

[5] We decided to present varying neutral stimuli instead of one constant neutral word in order to keep the neutral condition consistent with the congruent and incongruent conditions in this respect. However, regarding the response competition setting, previous results showed that the use of a single versus varying neutral primes has little relevance for the influence of primes and its division into costs and benefits (Mahr & Wentura, 2014).

assignment counterbalanced). Participants were informed that the auditory words were not predictive to the target.

To start each trial, participants pressed the space bar. On each trial, following a 1,000 ms blank (black) screen, a white fixation cross appeared for 500 ms, followed by a blank screen without sound for approx. 390 ms (there was a slight jitter due to the subsequent audio file loading). The auditory prime commenced during the presentation of the blank screen and was followed by the target screen with a SOA of 90 ms. The target screen was presented until a response was given (see *Figure 1*).

Participants completed two practice phases. After being presented with all color circles (i.e., the target and distractor circles), participants first completed 24 single-item trials, in which they simply had to indicate whether a circle had a target color or not. If necessary, this phase was repeated until participants reached at least 90% accuracy. Then, they proceeded to a further practice block of 30 trials, in which the main experiment was practiced (i.e., a sub-sample of the full design was shown, with the different factors balanced). Feedback was given on each trial.

Afterwards, participants completed two experimental blocks of 120 trials each, composed of 24 neutral, 24 congruent, and 72 incongruent trials, randomly intermixed. The trial list was balanced with regard to the four colors (i.e., each of the 20 possible prime-target combinations was presented six times; two times in each set size condition: one time with an "up" gap, one time with a "down" gap). Within the "up" and "down" gap conditions, it was randomized whether the -45°, 0°, or 45° position of the gap was used in a given trial (see *Materials*). The gaps of the distractors were randomly drawn from the six possible realizations. Four warm-up-trials preceded each block (not included in the analyses). The experiment took approximately 25 minutes.

Results

All effects referred to as statistically significant throughout the text are associated with *p* values below .05, two-tailed, except for the congruency effect that is tested one-tailed (i.e., because of the expected positive congruency effect on RTs, thus, faster responses in trials with congruent priming compared to trials with incongruent priming; see *Preregistration*). We report all measures, manipulations, and exclusions in all experiments of the study.

Error rate was 8.3% (*SD* = 10.1%). Trials with RTs below 150 ms or RTs that were greater than 1.5 interquartile ranges above the third quartile with respect to the individual distribution of RTs in the respective set size condition were discarded (Tukey, 1977); this led to exclusion of 5.0% of trials. Mean RTs and error rates are reported in *Table 1* (see also *Figure 2*).

Repeated Measures MANOVA

To test our hypotheses, we conducted a 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) MANOVA for repeated measures (e.g., Dien & Santuzzi, 2005; O'Brien & Kaiser, 1985). As expected, the analysis revealed a significant main effect of priming, $F(2,43) = 15.45$, $p < .001$, $\eta_p^2 = .418$. The a priori contrast of main interest, congruent versus incongruent trials, revealed a significant difference, $F(1,44) = 30.87$, $p < .001$, $\eta_p^2 = .412$. As expected, responses in congruent trials were faster than responses in incongruent trials (see *Table 1* and *Figure 2*). The main effect of set size was significant as well, $F(2,43) = 120.38$, $p < .001$, $\eta_p^2 = .848$. As expected, this effect was due to a linear trend, $F(1,44) = 214.04$, $p < .001$, $\eta_p^2 = .829$, with slower responses for larger displays. The quadratic trend was not significant, $F(1,44) = 2.47$, $p = .123$, $\eta_p^2 = .053$.

Most importantly, there was a significant interaction of priming and set size, $F(4,41) = 7.21$, $p < .001$, $\eta_p^2 = .413$. This interaction is exclusively due to the interaction of the congruent versus incongruent contrast of priming and the linear trend of set size, $F(1,44) = 19.18$, $p < .001$, $\eta_p^2 = .304$ (*F*s < 2.86, *p*s > .098, for the further interaction contrasts). As can be seen in Figure 2, the effect of set size is reduced in congruent trials compared to incongruent trials, providing evidence for our search efficiency hypothesis.

Although each priming effect for the three set sizes was significant, they increased in terms of mean RT-difference and effect size with increasing set size: for set size = 2, $t(44) = 1.91$, $p = .032$, $d_Z = 0.28$; for set size = 8, $t(44) = 3.93$, $p < .001$, $d_Z = 0.59$; for set size = 16, $t(44) = 5.14$, $p < .001$, $d_Z = 0.77$. This result already indicates that the increase in priming with set size is not simply a scaling phenomenon (i.e., that a larger RT-difference is an artefact of generally slower responses at larger set sizes). To completely rule out this possibility, however, we calculated priming differences as a proportion of base RT (i.e., RTs for neutral priming for a given set size). These relativized priming scores were 2.2% (*SE* = 1.2%), 6.2% (*SE* = 1.6%), and 12.9% (*SE* = 2.1%) for set sizes 2, 8, and 16, respectively. The linear increase was significant, $F(1,44) = 18.50$, $p < .001$, $\eta_p^2 = .296$. In addition, difference scores for benefits (neutral – congruent RT) and costs (incongruent – neutral RT) are reported for all set sizes in *Table 1.* Descriptively, the benefits outweigh the costs for smaller set sizes, while they were more balanced at the largest set size. Indeed, benefits were significant for each set size conditions, $t(44) = 2.05$, $p = .046$, $d_Z = 0.31$, for set size 2; $t(44) = 3.71$, $p < .001$, $d_Z = 0.55$, for set size 8; and $t(44) = 2.75$, $p = .009$, $d_Z = 0.41$, for set size 16. In contrast, costs were significant only for the largest set size, $t(44) = 4.93$, $p < .001$, $d_Z = 0.73$; whereas

$t(44) = -0.57$, $p = .570$, $d_Z = 0.09$; and $t(44) = 0.42$, $p = .678$, $d_Z = 0.06$, for set sizes 2 and 8, respectively.

On error rates, a 3 (prime-target relation) × 3 (set size) MANOVA for repeated measures revealed no significant overall effects, $F(2,43) = 2.38$, $p = .105$, $\eta_p^2 = .100$ for priming, $F < 1$ for set size main effect and the interaction. There were numerically more errors in the incongruent condition compared to the congruent condition (see *Table 1*), mirroring the RT priming effect, but this was non-significant, $F(1,44) = 3.35$, $p = .074$, $\eta_p^2 = .071$. Numerically, the priming effects also increased from low set size to large set size; the corresponding interaction contrast (i.e., congruent vs. incongruent × linear trend) was, however, not significant, $F(1,44) = 1.53$, $p = .223$, $\eta_p^2 = .034$ ($F < 1$ for congruent vs. incongruent × quadratic trend). In sum, there were no indications of a speed-accuracy tradeoff with regard to the RT effects.

Linear Mixed Model Analyses

In order to test the influence of priming on the search slopes, we used LMM analysis (using lmerTest package, Kuznetsova, et al., 2017; Bates, et al., 2015): We analyzed RTs as a function of priming condition (coded as +1 congruent, -1 incongruent), set size (i.e., 2 / 8 / 16), and their interaction. We fitted the maximal model including the full variance-covariance structure of random effects (see e.g., Barr et al., 2013). Thus, we allowed random intercepts and slopes for participants. Set size had a weight of $b = 28.3$ ms ($SE = 1.7$ ms), $t(43.71) = 16.65$, $p < .001$, indicating that with each additional element in the display, search RTs increased by 28 ms. The interaction of set size and priming was significant as well, $b = -4.9$ ms ($SE = 1.2$ ms),

$t(48.36) = 4.17$, $p < .001$. Slopes were flatter for congruent ($b = 23.4$ ms) than for incongruent pairs ($b = 33.2$ ms).[6]

Discussion

Experiment 1 provided direct evidence that cross-modal semantic priming influences search efficiency, indicating that auditory primes can guide visual search processes. We found not only a congruency effect by auditory primes (i.e., faster RTs in congruent as compared with incongruent trials), but also that this congruency effect increased with increasing set size. Accordingly, the corresponding LMM analyses showed that the search slope was significantly flatter in case of congruent priming compared to incongruent priming. Thus, these results have important implications regarding the involved processes: The evidence that semantically congruent auditory primes can increase the efficiency of visual search indicates that the processes involved are not limited to facilitation of target encoding, which is assumed to be the dominant process of semantic priming without spatial uncertainty of the target (e.g., McNamara, 2005, 2013). Instead, this pattern of results is consistent with the assumption that auditory primes facilitate the guidance of attentional selection in favor of the primed target, tentatively, by increasing activity on the priority map for the features of that target by sensory gain modulation (see, e.g., Yu et al., 2023; Wolfe, 2021).

The cost and benefit calculation yielded a dominance of benefits in smaller set sizes (i.e., thereby replicating the results of Mahr & Wentura, 2014), while benefits and costs were balanced at the largest set size (see *Table 1*). Thus, facilitation appears to be

[6] A more complex analysis including the neutral trials (with two contrast codes for priming and therefore two interaction terms) yielded almost the same result for the interaction contrast of congruent vs. incongruent with set size, $b = -4.9$ ms ($SE = 1.2$ ms), $t(47.9) = 4.15$, $p < .001$, and, as expected, a non-significant result for the interaction contrast of neutral vs. congruent/incongruent with set size, $b = -0.8$ ms ($SE = 1.1$ ms), $t(133.2) = 0.71$, $p = .475$.

the main component of the priming effect for less difficult search, whereas for more complex searches, in which bottom-up target information decreases, prioritization of the incongruently primed target is associated with increasing costs. Cost-benefit analyses should be taken, however, with a grain of salt, because the choice of type of a neutral baseline condition may affect the interpretation of these results (see Mahr & Wentura, 2014, for a detailed discussion).

In order to manipulate the semantic congruency between primes and targets, we presented a target set of four potential targets. Thus, the task required participants to maintain four possible target items in working memory and use these representations to guide visual attention. Although real-life search behavior is often characterized by searching for multiple potential targets, visual search experiments most often examine search processes for a single target, requiring participants to use a single attentional template. Recent research has shown that there is a potential cost to using multiple attentional templates compared to single-target search. Specifically, there is evidence of costs to both maintaining the mental representations of multiple search targets and to selecting multiple targets; furthermore, these costs increase with increasing target set (e.g., Ort & Olivers, 2020). Importantly, searching for multiple potential targets in the present experiment may have contributed not only to a less efficient visual search in general, but also to the occurrence of the congruency effect by auditory primes. This is because an increase in cognitive load, typically operationalized as working memory load, can increase the influence of task-irrelevant stimuli in visual search tasks by impeding inhibitory processes (e.g., Burnham, 2010; Lavie, 2010; Murphy et al., 2016). As indicated by abundant empirical evidence (for reviews, see Lavie, 2010; Murphy et al., 2016), the increased demand on executive control processes under high (versus low) working memory load can lead to reduced inhibition of the influence of goal-irrelevant items on task-relevant processes. Because in our task the auditory primes were not predictive of the target identity, moreover, participants were informed about this, prime information was completely task-irrelevant (i.e., irrelevant regarding the current task goals). Hence, it can be argued that the working memory load of maintaining a relatively large target set might have potentiated the influence of auditory primes by hindering the suppression of the task-irrelevant prime information. In line with this argumentation, previous research investigating overt attentional orienting by using eye-tracking (Belke et al., 2008) showed that working memory load (i.e., maintaining task-unrelated items in working memory) during visual search can lead to increased interference from visual distractors that are semantically related to the target. It should be noted, however, that this relation was found in terms of post-search processes, as reflected by increased viewing time of the semantically related distractors and also of the target under high (versus no) cognitive load. The authors interpreted this result as cognitive load influencing the process of rejecting or accepting an item as a target. In contrast, working memory load did not affect the initial fixations to the semantically related distractors.

Given this argumentation, it is important to test the observed cross-modal facilitation without the cognitive load of maintaining four items in working memory simultaneously. Therefore, we reduced the set of potential targets to the smallest set that allows for the manipulation of semantic congruency in the present setting, that is, to two potential targets. Using simple stimuli, that is, target colors, there is evidence that dual-target search may allow attention to be directed by multiple attentional templates without considerable costs (Kerzel & Witzel, 2019).

In the following, we pursued two goals: First, as outlined above, we aimed to test whether the influence of auditory primes on visual search processes holds when working memory load is reduced (Experiment 2). Second, we aimed to test if this search efficiency hypothesis holds for a setting that is more relevant to real-world search behavior, thus, when attentional sets contain perceptually and semantically complex object representations (Experiment 3). Because the data collections of these experiments were conducted during the COVID-19 pandemic, recruiting a sufficient sample size was not realistic using the approach that is typical of cognitive psychology, thus, testing a moderate to high number of participants (e.g., $N = 45$; see the power calculation of Experiment 2) in a moderate number of trials per participant and condition (i.e., typically 10-100; Rouder & Haaf, 2018; in the present Experiment 1, there were 16 congruent and 48 incongruent trials in each set size condition). Since power is derived from both the number of participants and the number of repeated measurements per participant and condition, under certain assumptions, the number of observations can be distributed among participants and trials without compromising power (e.g., Rouder & Haaf, 2018; see Brysbaert & Stevens, 2018, for simulation studies; see also Smith and Little, 2018, for simulations studies using a small-*N* approach to test individual-level effects with high power). We decided to compensate for the reduction in the number of participants by increasing the number of trials that was presented to each participant, and to test the feasibility of this approach beforehand. Therefore, we aimed to replicate the results of Experiment 1 using a small-*N* design (regarding typical semantic priming, see Miller, 2023, that indicates the feasibility of this approach). Hence, we conducted an exact replication of Experiment 1, except that we used a small-*N* strategy, that is, we collected a comparable total number of observations per condition as in Experiment 1, but with $N = 6$ participants. We report the replication of Experiment 1 (Experiment 1R in the remainder) as well as a more detailed discussion on this topic in the *Appendix*. In a nutshell, this experiment replicated the facilitation of search efficiency by congruent priming using a small-*N* approach. The most central result, the interaction between congruent versus incongruent priming and set size (linear contrast) was associated with a large effect size (see *Appendix*).

Experiment 2

To exclude the possibility that the observed effects are restricted to high working memory load, the relevant target set was reduced to two items in Experiment 2; in each task block, the two relevant target colors were chosen from the set of four target color (i.e., blue, green, yellow, red). In six experimental blocks, all possible pairs of these four colors were presented in random order. We again varied display size. Materials and task were identical to those used in Experiment 1 (and in its small-*N* replication).

Overall, we expected a significant priming effect, thus, faster responses in trials with congruent priming compared to trials with incongruent priming across display size conditions. Furthermore, we tested whether RTs overall would increase with increasing display size when using a reduced target set (i.e., indication of serial search processes); and, importantly, if priming effects would increase with increasing display size, indicating more efficient search with congruent priming.

In addition, since Experiment 2 had an identical design to Experiment 1R (see *Appendix*), except for the manipulation of the potential target set, and both used a small-*N* approach, we examined possible differences between these experiments in terms of search efficiency and the influence of primes on RTs and search slopes. To address these questions, we added combined analyses across Experiment 1R and the present

Experiment 2, including experiment (i.e., using 4 vs. 2 potential targets, respectively) as a between-participants factor to our preregistered analyses.

Method

The experiment was preregistered (https://aspredicted.org/FLD_867). We followed our preregistered procedure with the following main exception (for minor points, see *Table A2*): As preregistration took place *before* and data collection took place *during* the COVID-19 pandemic, in order to reach the number of observations specified in the preregistration, we compensated for a reduction in the number of participants by increased number of trials per participant (for more details, see *Appendix* and the *Participants* section).

Participants

For power calculations, we oriented ourselves on the results of Experiment 1: The overall priming effect (i.e., congruent versus incongruent comparison as the contrast of main interest) was associated with $d_Z = 0.83$; and the interaction between priming (congruent versus incongruent) and set size (linear contrast) was associated with $d_Z = 0.65$. The latter interaction contrast tests for the central hypothesis: an enhancement of the priming effect with increasing display size. Since we changed the task (i.e., reduced target set), we reduced the expected effect size for the interaction to $d_Z = 0.50$. To detect an effect of $d_Z = 0.5$ with $\beta = .95$ ($\alpha = .05$, one-tailed), a sample size of at least $N = 45$ is needed with one run of the experimental procedure per participant (see *Preregistration*).

Such a proceeding (i.e., $N = 45$ with 432 trials per participant: the experiment featured now six blocks in order to present each possible combinations of the target colors; see *Procedure*) would have yielded a total number of observations of 19440.

We decided to adopt a small-*N* approach (i.e., to compensate for the reduction in the number of participants by increasing the number of trials per participant), hence, we decided for the sample size of $N = 12$ participants in a testing procedure that comprised 3456 trials per participant. The testing procedure was realized in two testing sessions on two consecutive days. Each testing session consisted of four runs of the complete experimental procedure (i.e., 432 trials each). The procedure resulted in a total number of 41472 observations for the sample.

The final sample size was $N = 13$, as due to a misunderstanding of experimenters when carrying out the data collection, two participants completed only one half of the testing sessions (i.e., both participants worked through four runs of the experimental procedure on one day each). These two data sets were included in the final sample; therefore, the total number of observations was the same as the a priori planned total number of observations for $N = 12$. (Excluding these participants from the analyses yielded essentially the same results as reported below.) Participants were 13 students (7 women, 6 men) from Saarland University, who took part in the experiment for 50 Euro. The median age was 23 years (ranged 20 to 30). All participants were native speakers of

German and had self-reported normal hearing and normal or corrected-to-normal vision. Data from none of the participants met the preregistered criteria for outlier exclusion.

Design and Statistical Analyses

Experiment 2 followed a 3 (prime-target relation: congruent, incongruent, neutral) × 3 (set size: 2, 8, 16 stimuli) design, with all factors manipulated within participants. Technically, the congruency factor was realized by a 3 (auditory prime type: color 1, color 2, neutral) × 2 (visual target type: color 1, color 2) design, in which color 1 and color 2 denotes the two target colors of the current task block. Primes and targets were thus uncorrelated, resulting in auditory priming without contingency.

In line with Experiment 1 and the preregistration, we conducted a 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) MANOVA for repeated measures, and we a priori focused on the contrast between congruent and incongruent conditions regarding priming. We expected a linear relationship between set size and search RTs (i.e., the linear trend of set size was of main interest). Furthermore, to test the influence of priming on search slopes more directly, as preregistered, we report linear mixed model (LMM) analyses: We again analyzed RTs as a function of priming condition (i.e., congruent / incongruent), set size (i.e., 2 / 8 / 16), and their interaction.

In addition to the preregistered analyses, we compared the present experiment with Experiment 1R in terms of the main analyses. Specifically, first, we conducted a 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) × 2 (experiment: Experiment 1R vs. Experiment 2) MANOVA for repeated measures with prime-target relation and set size as within-participants factors and experiment as between-participants factor. Second, we added experiment (i.e., corresponding to a target set of 2 vs. 4) to our main LMM analysis testing the effect of congruent versus incongruent priming on search slopes.

Material and Procedure

Materials were the same as in Experiment 1. The experimental procedure was identical to Experiment 1, with the following exceptions: The relevant target set size was reduced to two items; in each task block (i.e., 72 trials), the two relevant target colors were chosen from the set of four target color (i.e., blue, green, yellow, red). As in Experiment 1R with a small-*N* approach (see *Appendix*), each participant was tested in two testing sessions that took place on two consecutive days. Each testing session consisted of four runs of the experimental procedure (i.e., 432 trials each). All possible pairings of the four target colors used in Experiment 1 were presented in random order in the six blocks of the experimental procedure. As in Experiment 1, participants completed two practice phases: After being presented with all color circles, participants first indicated whether a circle had a target color or not. Then, they proceeded to a further practice block of 36 trials, in which the experimental task was practiced. This practice block featured a random pair of the four possible target colors as targets. Afterwards, participants completed six experimental blocks of 72 trials each, composed of 24 neutral, 24 congruent, and 24 incongruent trials, randomly intermixed. Each of the 6 possible prime-target combinations was presented twelve times: four times in each set size condition, and twice with an "up" gap, twice with a "down" gap. All other details of the procedure were the same as in Experiment 1. One experimental procedure took approximately 35 minutes, one testing session (i.e., four runs of the experimental procedure) took about three hours including the mandatory short breaks between the experimental runs (i.e., approx. 5 minute each) and a 30 minute break after the first two runs.

Results

Error rate was 3.1% (*SD* = 1.9%). Trials with RTs that were greater than 1.5 interquartile ranges above the third quartile with respect to the individual distribution of RTs in each set size condition were discarded (Tukey, 1977); this led to exclusion of 4.5% of trials. There were no RTs below 150 ms (i.e., preparatory responses). Mean RTs and error rates are reported in *Table 2* (see also *Figure 3*).

Repeated Measures MANOVA

The 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) MANOVA for repeated measures (e.g., Dien & Santuzzi, 2005; O'Brien & Kaiser, 1985) revealed a significant main effect of prime-target relation, $F(2,11) = 5.97$, $p = .018$, $\eta_p^2 = .521$. Importantly, the congruency effect (i.e., congruent vs. incongruent trials) was significant, $F(1,12) = 12.98$, $p = .004$, $\eta_p^2 = .520$. Responses in congruent trials were faster than responses in incongruent trials (see *Table 2* and *Figure 3*). The main effect of set size was significant, $F(2,11) = 138.93$, $p < .001$, $\eta_p^2 = .962$. As expected, search RTs were longer with larger search displays, indicating serial search processes, $F(1,12) = 254.60$, $p < .001$, $\eta_p^2 = .955$, for the linear trend. ($F[1,12] = 4.57$, $p = .054$, $\eta_p^2 = .276$, for the quadratic trend of set size.) The overall interaction between prime-target relation, including the neutral condition, and set size was not significant, $F(4,9) = 1.78$, $p = .217$, $\eta_p^2 = .442$. Importantly, the interaction between congruent versus incongruent priming and the linear trend of set size was significant, $F(1,12) = 7.49$, $p = .018$, $\eta_p^2 = .384$. (For all other interaction contrasts, *F*s < 2.16, *p*s >.167).

Thus, the influence of set size on search RTs was reduced in congruent trials compared to incongruent trials (see *Figure 3*).

In each set size, the congruency effects were significant, for set size 2, $t(12) = 3.13$, $p = .004$, $d_Z = 0.87$; for set size 8, $t(12) = 3.89$, $p = .001$, $d_Z = 1.08$; for set size 16, $t(12) = 3.12$, $p = .004$, $d_Z = 0.87$. Furthermore, priming differences (i.e., difference between congruent vs. incongruent trials) as a proportion of base RT (i.e., RTs for neutral priming for a given set size) increased with increasing set size, $F(1,12) = 6.98$, $p = .021$, $\eta_p^2 = .368$; these relative priming differences were 1.2% (*SE* = 0.4%); 2.1% (*SE* = 0.5%); and 4.1% (*SE* = 1.3%) for set size 2, 8, and 16, respectively.

Difference scores for benefits (see *Table 2*) were significant for set sizes 2 and 16, $t(12) = 2.58$, $p = .024$, $d_Z = 0.72$; and $t(12) = 2.24$, $p = .045$, $d_Z = 0.62$, respectively; while it missed significance for set size 8, $t(12) = 1.66$, $p = .124$, $d_Z = 0.46$. Costs were not significant for the lowest set size, $t(12) = 1.52$, $p = .153$, $d_Z = 0.42$; while significant for set size 8 and 16, $t(12) = 2.50$, $p = .028$, $d_Z = 0.69$, and $t(12) = 2.78$, $p = .017$, $d_Z = 0.77$, respectively.

For the error rates, the corresponding analysis revealed no significant effects, *F*s < 1.28, *p*s > .345. The congruency effect was also not significant, $F(1,12) = 0.02$, $p = .901$, $\eta_p^2 = .001$; nor was its interaction with the linear trend of set size, $F(1,12) = 2.20$, $p = .164$, $\eta_p^2 = .155$ (see also *Table 2*).

Linear Mixed Model Analyses

To corroborate the MANOVA analyses, we tested the search efficiency hypothesis using LMM-approach: RTs (of correct responses; outliers discarded; see above) were analyzed as a function of the priming condition (coded as +1 congruent, -1 incongruent), set size (i.e., 2 / 8 /16), and their interaction by using LMM analysis (lmerTest package, Kuznetsova et al., 2017; Bates, et al., 2015). We again fitted the maximal model (e.g., Barr et al., 2013). In this random slopes model, set size had a weight of b = 12.5 ms (SE = 0.8 ms), $t(11.99) = 15.47$, $p < .001$, indicating that with each additional element, search RTs increased by 12.5 ms. The interaction of set size and priming was significant as well, b = -0.8 ms (SE = 0.3 ms), $t(10.88) = 2.65$, p = .023. Slopes were thus flatter for congruent (b = 11.7 ms) than for incongruent pairs (b = 13.3 ms).[7]

Across-Experiments Analyses

Repeated Measures MANOVA. We conducted a 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) × 2 (experiment: Experiment 1R vs. Experiment 2) MANOVA for repeated measures on RTs with prime-target relation and set size as within-participants factors and experiment as between-participants factor. The factor Experiment 1R versus Experiment 2 correspond to a comparison of target set of 4 versus 2, respectively. The results of this analysis are reported in *Table 3*. Overall, the results on the combined data from the two experiments corroborate the results of the individual experiments (see *Effects across Exp. 1R and 2* in *Table 3*; as an exception, the interaction between congruent versus incongruent priming and the quadratic trend of set size also reached significance in this analysis, indicating that the congruency effect at medium set size was not exactly the average of the congruency effects at small and large set sizes). In general, RTs did not differ between experiments ($F[1,17] = 2.77$, $p = .114$, $\eta_p^2 = .140$). Of particular importance,

[7] The more complex analysis including the neutral trials (with two contrast codes for priming and two interaction terms), again, supported the pattern of results reported in the main text: The interaction contrast of congruent vs. incongruent priming with set size was significant, b = -0.8 ms (SE = 0.3 ms), $t(10.79) = 2.61$, $p = .025$. The interaction contrast of neutral vs. congruent/incongruent priming with set size was non-significant, b = 0.0 ms (SE = 0.2 ms), $t(13.32) = 0.02$, $p = .983$.

we report in *Table 3* the comparisons between Experiments 1R and 2 regarding each effect (i.e., the interactions including the experiment factor). Looking at the most important comparisons, the experiments did not differ significantly with respect to priming, the interaction of congruent versus incongruent priming and experiment also missed significance. Search was more efficient in Experiment 2 (see *Table A1* and *Table 2*), as indicated by the significant interaction of the linear trend of set size and experiment. Importantly, the interaction between congruent versus incongruent priming and the linear trend of set size was significantly moderated by the experiment. Thus, the effect of set size was reduced more by congruent (versus incongruent) priming in Experiment 1R than in Experiment 2 (see *Table A1* and *Table 2*).

Linear Mixed Model Analyses. To address between-experiment differences regarding search slopes more directly, we included experiment as a further predictor in the main LMM analysis reported above. Thus, RTs were analyzed as a function of the priming condition (coded as +1 congruent; -1 incongruent), set size (i.e., 2 / 8 /16), and experiment (Experiment 1R was coded as -1; Experiment 2 was coded as +1), including all two-way and the three-way interaction terms. Across the two experiments, set size had a weight of b = 15.0 ms (SE = 1.0 ms), $t(17.15) = 15.32$, $p < .001$. The interaction of set size and experiment was significant, b = -2.4 ms (SE = 1.0 ms), $t(17.15) = -2.46$, $p = .025$. RTs increased 4.8 ms less for each additional item in the display in Experiment 2 than in Experiment 1R, indicating more efficient search. Across experiments, the interaction of set size and priming was significant, b = -1.4 ms (SE = 0.3 ms), $t(21.28) = -4.75$, $p < .001$. This interaction was moderated by experiment, b = 0.6 ms (SE = 0.3 ms), $t(21.28) = 2.13$, $p = .045$, indicating greater influence of priming on search slopes in Experiment 1R compared to Experiment 2.

Discussion

Experiment 2 successfully replicated the essential findings of Experiment 1: Search slopes were significantly flatter when the primes were congruent to the target compared to incongruent trials. Furthermore, as expected, Experiment 2 with two potential targets was indeed associated with more efficient visual search than Experiment 1R, which also used a comparable small-*N* approach, and required participants to keep four attentional templates active in working memory. While in Experiment 2 the RT-increase was 12.5 ms per additional item, in Experiment 1R, search RTs increased by 17.4 ms with each additional item. Regarding the influence of primes, priming differences (also as a proportion of base RT) were numerically larger when using a larger target set (see Experiment 1 and Experiment 1R) than in Experiment 2. However, the comparison of the priming differences between Experiment 1R and Experiment 2 missed significance. By using a smaller target set in the present experiment, the influence of primes on search efficiency was significantly reduced compared to Experiment 1R. Thus, on the one hand the result fit to the critical argument that the pattern found in Experiment 1 might be partly due to the increased demand on executive control processes by maintaining a larger target set.

On the other hand, however, we were able to replicate the decisive pattern (increased priming with increased set size) with the small target set. The priming effects were significant in each set size condition in Experiment 2, and the priming differences increased with increasing set size (in contrast to Experiment 1, the effect size of the priming differences did not increase from set size 8 to 16, but relative priming differences as a proportion of base RT increased with increasing set size). These results were corroborated by the LMM analysis, indicating more efficient search with congruent compared to incongruent priming. Thus, even with a reduced target set and thereby more efficient search, there was a significant influence of primes, moreover, this influence increased with increasing set size, albeit to a lesser extent than when using a larger target set.

To conclude, while it is possible that the working memory load of maintaining a relatively large target set contributed to the influence of non-predictive primes, the search efficiency hypothesis of cross-modal priming was supported even with a markedly reduced target set. Thereby, Experiment 2 showed clear evidence for the influence of cross-modal semantic priming on visual search efficiency with a reduced working memory load. Taken together, while the results of three experiments (i.e., including the Experiment 1R, see *Appendix*) shows that the guidance of visual attention by auditory primes can be involved in cross-modal semantic congruency effects, it remains to be investigated whether this process is restricted to simple perceptual features or represents a general mechanism of audio-visual information processing. Therefore, Experiment 3 tested whether the search efficiency hypothesis also holds for more complex visual search environments in which attentional templates contain perceptually and semantically complex object representations.

Experiment 3

While in Experiments 1 and 2 a simple perceptual feature defined the potential targets, in Experiment 3 we used complex warning icons from the automotive context as potential targets (see also Mahr & Wentura, 2018). Target symbols (e.g., ambulance, tractor) were defined by their semantic meaning, while they were perceptually similar to each other and to the distractors (e.g., train, bicycle, see *Figure 4*). Accordingly, Experiment 3 had two important novel aspects: First, participants had to search for complex objects defined by multiple perceptual and semantic features. In this regard, the primes did not directly address a perceptual feature by naming a potential target color, but they contained object-level semantic information. Therefore, we were able to test whether our hypothesis about facilitation of search efficiency involved in cross-modal priming generalizes to complex object representations. Second, as we presented warning symbols relevant to an automotive context, we were able to test the search efficiency hypothesis in a setting where it is of utmost practical importance. If semantically related auditory primes not only facilitate the processing of visual objects when they are attended, but also influence attentional guidance toward them, this would mean that auditory semantic cues could be used as efficient warning signals in critical traffic situations where the relevant objects are initially unattended. To apply the hypothesized process to a driving scenario, one might ask the question: would, for example, a spoken warning "Children!" increase attentional priority and thereby relative salience of initially unattended children on the road?

Specifically, Experiment 3 presented auditory primes (i.e., spoken denotations of automotive symbols) that could be semantically congruent, incongruent, or neutral to the automotive symbols being searched. In line with Experiments 1 and 2, auditory primes were again non-predictive with regard to the target identity (i.e., primes naming a potential target and targets were uncorrelated). The target symbols were presented in context with irrelevant symbols, and importantly, we again varied the size of these search displays to test the influence of priming on visual search slopes. Furthermore, as in Experiment 2, we again wanted to use low working memory load; thus, the relevant target set was two items. We expected that RTs overall would increase with increasing display size, as an indication of serial search processes. If the facilitation of search efficiency by auditory primes represents a semantic object-based interaction, we would expect not only a significant priming effect, but also that priming effects would increase with increasing display size. As the design of Experiment 3 was dominantly based on the design of Experiment 2, with the main difference being the complexity of the stimulus material (i.e., corresponding to feature- versus object-based search)[8], we also examined possible differences between these experiments in terms of search efficiency and the influence of primes on RTs and search slopes.

Method

The experiment was preregistered at https://aspredicted.org/blind.php?x=vh99jg.

Participants

For power calculations, we oriented ourselves on the results of Experiment 1[9]: The effect of main interest, the interaction between congruent versus incongruent priming and the linear contrast of set size was associated with $d_Z = 0.65$. As we changed the task, used a target set with two targets, and presented more complex stimulus material, we reduced the expected effect size to $d_Z = 0.50$ (i.e., a medium effect according to Cohen, 1988). To detect an effect of this size with $\beta = .95$ ($\alpha = .05$), a sample size of at least $N = 54$ is needed, assuming 240 observations per participant as in Experiment 1. We again used a small-*N* approach: We compensated for a reduction in the number of participants with increased number of trials per participant (see *Appendix*

[8] Of course, there were other minor differences between Experiment 2 and Experiment 3 that were related to the complexity of the material (oriented on Experiment 1 of Mahr and Wentura, 2018, using automotive symbols). These minor differences were: (1) the primes in Experiment 3 had a longer duration of 350 ms (vs. 120 ms in Experiment 2); (2) these more complex primes had a lower degree of compression (i.e., compression to 50% vs. 30% of the original duration in Experiments 3 and 2, respectively); (3) Experiment 3 had a slightly longer SOA of 100 ms (vs. 90 ms in Experiment 2); furthermore, (4) Experiment 2 used four possible neutral primes (see also Footnote 5), whereas Experiment 3 used one invariant neutral prime.

[9] We oriented ourselves on the results of Experiment 1 because data collection of Experiment 2 had not been completed at the time of the preregistration of Experiment 3.

and Experiment 2). With $N = 54$ and 240 observations per participant, the total number of observations would have been 12960. We a priori decided to recruit $N = 12$ participants, with each participant taking part in two sessions, with a total of six runs of the experimental procedure of 432 trials each (i.e., as in Experiment 2, one experimental procedure comprised six blocks due to the balancing of the two possible targets out of the pool of four). The procedure resulted in a total number of 31104 observations in the sample, and 3456 observations per experimental condition (see *Preregistration*).

Participants were twelve students (10 women, 2 men) from Saarland University, who took part in the experiment in exchange for 55 Euro. The median age was 24 years (ranged 18 to 27). All were native speakers of German and had self-reported normal hearing and normal or corrected-to-normal vision. Data from none of the participants met the preregistered criteria for outlier exclusion.

Design and Statistical Analyses

The study featured a 3 (prime-target relation: congruent, incongruent, neutral) × 3 (set size: 2, 8, 16) repeated measures design. The congruency factor was technically realized in a 3 (prime: symbol 1, symbol 2, neutral) × 2 (target: symbol 1, symbol 2) sub-design, in which symbol 1-2 denotes the target set of the actual block. Thus, primes and targets were uncorrelated, resulting in auditory priming without contingency (i.e., 1/3 congruent trials).

As in Experiments 1 and 2, we first report the results of a 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) MANOVA for repeated measures. As in the previous experiments, the contrast between congruent and incongruent conditions regarding priming and the linear trend of set size were of main interest. Then, as preregistered, we report the LMM analysis testing the influence of priming on search slopes. Please note that our preregistration focused solely on LMM analyses to test the corresponding hypotheses; we first report the MANOVA analysis to increase comparability with the results of Experiments 1 and 2 (see *Table A2*).

In addition to our preregistered analyses, we compared the present experiment with Experiment 2 to test for possible differences regarding the complexity of the stimulus material. Specifically, first, we conducted a 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) × 2 (experiment: Experiment 2 vs. 3) MANOVA for repeated measures with prime-target relation and set size as within-participants factors and experiment as between-participants factor. Second, we added experiment to our main LMM analysis (i.e., testing the effect of congruent versus incongruent priming on search slopes). The factor Experiment 2 versus 3 dominantly correspond to feature- versus object-based search, respectively, both with two potential targets.

Material

Each target display contained two, eight, or sixteen symbols presented in random locations of a 4 × 4 matrix (18.5 × 18.5 cm, approx. 18.5 × 18.5 ° visual angle; see *Figure 4*). Each symbol was drawn in black to avoid orientating on color features. Symbols were centered on a white square background with the size of approx. 3.6 × 3.6 cm (i.e., 73 x 73 pixels on the screen). Each search display contained one of the two possible target symbols of the block (throughout the experiment, we presented four possible target symbols; i.e., traffic light, children, ambulance, tractor; adapted from Mahr and Wentura, 2018, with minor adjustments to allow left and right alignments). Furthermore, depending on the display size condition, each display featured one, seven, or fifteen distractors. Distractors were randomly selected from a set of 16 symbols representing relevant objects in the driving context (e.g. roadwork, bicycle, barrier, gas station; see *Figure A2*). Distractor symbols were clearly distinguishable from the target symbols. For each target and distractor symbol, there was a version that faced to the left, and one that faced to the right as the task-relevant feature (see *Figure A2*).

Four German two-syllable spoken denotations of the target symbols (“traffic light”/”Ampel”; “children”/”Kinder”; “ambulance”/”Notarzt”; “tractor”/”Traktor”) and the word “object” (“Objekt”; i.e., the neutral condition) were presented as auditory primes (adapted from Mahr & Wentura, 2018). Auditory primes were presented time-compressed to 50% of their original duration, i.e., to 350 ms duration (see Mahr and Wentura, 2018), to allow for presentation with a brief SOA. The sounds were presented with closed over-ear headphones (AKG K511) at a sound pressure level of 65 to 71 dB SPL.

Procedure

As in Experiment 1R and in Experiment 2, each participant was tested in two testing sessions that took place on two days. Each testing session consisted of three runs of the complete experimental procedure (i.e., 432 trials each, divided into six task blocks). Between the runs of the experimental procedure, participants were instructed to take a break of at least twenty minutes. In each task block, the two relevant target symbols were selected from a set of four possible target symbols. All possible pairings of the four target symbols were presented in random order in the six blocks of the experimental procedure.

Experimental procedure followed the procedure of Experiment 2, with the following exceptions: (1) the present experiment used automotive symbols presented in grayscale as targets and distractors instead of color feature (and accordingly, spoken denotations of the target symbols as primes); (2) we changed the task to categorization of the target symbol's orientation as facing to the left or to the right by pressing an assigned key (keys "D" and "L" on a standard QWERTZ keyboard for left and right alignment, respectively); (3) prime duration and prime-target SOA were identical to the Stroop-like experiment of Mahr and Wentura (2018; Experiment 1) using automotive symbols. On each trial, following a 1,000 ms blank white screen, a black fixation cross appeared for 500 ms. Thereafter, a blank screen was presented without sound for 250 ms. The 350 ms long auditory presentation started during the presentation of this blank screen and was followed by the target screen with a SOA of 100 ms (see *Figure 4*, for the illustration of a trial).

Results

Error trials were excluded from the RT-analyses (2.4% of the trials). Trials with RTs below 150 ms (i.e., preparatory responses) and with RTs that were greater than 1.5 interquartile ranges above the third quartile with respect to the individual distribution of RTs in each set size condition were discarded (Tukey, 1977); this led to exclusion of 6.9% of trials. Mean RTs and error rates are reported in *Table 3* (see also *Figure 5*).

Repeated Measures MANOVA

The 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) MANOVA for repeated measures revealed a significant main effect of prime-target relation, $F(2,10) = 26.89$, $p < .001$, $\eta_p^2 = .843$. As expected, there was a significant congruency effect, $F(1,11) = 50.33$, $p < .001$, $\eta_p^2 = .821$. The main effect of set size was also significant, $F(2,10) = 120.16$, $p < .001$, $\eta_p^2 = .960$, as was the linear trend of set size, $F(1,11) = 233.10$, $p < .001$, $\eta_p^2 = .955$, indicating serial search processes (while $F[1,11] = 3.26$, $p = .098$, $\eta_p^2 = .229$, for the quadratic trend). The interaction between prime-target relation and set size was also significant, $F(4,8) = 10.85$, $p = .003$, $\eta_p^2 = .844$. Importantly, there was a significant interaction between congruent versus incongruent priming and the linear trend of set size, $F(1,11) = 13.11$, $p = .004$, $\eta_p^2 = .544$ (Fs < 1, for the other interaction contrasts).

In each set size condition, the congruency effects were significant; ps < .001 (see *Table 4* for the priming effects). As expected, priming differences as a proportion of base RT (i.e., relative to the neutral condition of the corresponding set size) increased with increasing set size, $F(1,11) = 6.32$, $p = .029$, $\eta_p^2 = .365$; the relative priming differences were 4.5% (SE = 0.5%); 6.2% (SE = 0.9%); and 8.0% (SE = 1.6%) for set size 2, 8 and 16, respectively. Descriptively, the benefits and costs were balanced for each set sizes (see *Table 3*). Accordingly, benefits and costs were significant for each set size conditions; ps < .001 for benefits, while ts > 2.92, and ps < .014 for costs.

For the error rates, the corresponding analysis revealed a significant main effect of set size, $F(2,10) = 4.53$, $p = .040$, $\eta_p^2 = .475$; and a significant linear trend, $F(1,11) = 8.03$, $p = .016$, $\eta_p^2 = .422$. This effect reflects more errors with smaller set size conditions (see *Table 4*). Except for the effect of set size, there were no further significant effects, $F < 1$, for the main effect of priming; and $F(4,8) = 3.28$, $p = .072$, $\eta_p^2 = .621$, for the interaction. The congruency effect was also not significant, $F(1,11) = 0.53$, $p = .480$, $\eta_p^2 = .046$; similarly to the interaction between prime-target congruency and the linear trend of set size, $F(1,11) = 4.17$, $p = .066$, $\eta_p^2 = .275$. Thus, there was an indication for speed-accuracy-tradeoff regarding set size (i.e., set sizes with shorter RTs were also associated with somewhat more errors); however, accuracies were still very high in all set size conditions (i.e., 97.0%; 97.7%; and 98.1% for set size 2, 8, and 16, respectively). With respect to priming, there was no speed-accuracy-tradeoff.

Linear Mixed Model Analyses

RTs (of correct responses, outliers discarded) were analyzed as a function of the priming condition (coded +1 congruent, -1 incongruent), set size, and their interaction by using LMM analysis (lmerTest package, Kuznetsova et al., 2017, Bates, et al., 2015). To test the influence of priming on search slopes, we again fitted the maximal model. In this random slopes model, set size had a weight of b = 27.3 ms (SE = 1.8 ms), t(10.99) = 14.81, p < .001. Thus, search RTs increased with each additional element by 27.3 ms. As hypothesized, priming had an influence on the visual search slopes, b = -1.7 ms (SE = 0.5 ms), t(12.31) = -3.17, p = .008, with RTs increasing 3.4 ms less in congruent trials than in incongruent trials for each additional symbol in the display.[10]

Across-Experiments Analyses

Repeated Measures MANOVA. We conducted a 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) × 2 (experiment: Experiment 2 vs. Experiment 3) MANOVA for repeated measures on RTs with prime-target relation and set size as within-participants factors and experiment as between-participants factor. The results are reported in *Table 5*. In general, the comparisons without the experiment factor are consistent with the results of the individual experiments (see *Effects across Exp. 2 and 3* in *Table 5*). As expected, RTs were slower in Experiment 3 when complex stimuli were presented (F[1,23] = 10.88, p = .003, η_p^2 = .321). The comparisons between Experiments 2 and 3 regarding each effect (i.e., the interactions including the experiment factor) are also reported in *Table 5*. To reiterate the most important results,

[10] The more complex analysis including the neutral trials, again, supported the pattern of results reported in the main text: The interaction contrast of congruent vs. incongruent priming with set size was significant, b = -1.7 ms (SE = 0.5 ms), t(12.33) = -3.18, p = .008. The interaction contrast of neutral vs. congruent/incongruent priming with set size was non-significant, b = -0.2 ms (SE = 0.4 ms), t(18.62) = -0.41, p = .690.

there was a significant interaction between the congruent versus incongruent contrast of priming and experiment, priming differences were significantly larger in Experiment 3 than in Experiment 2 (see *Table 2* and *Table 3*). As expected, search processes were more efficient in Experiment 2 than in Experiment 3, as indicated by the significant interaction of the linear trend of set size and experiment. (The interaction between the quadratic trend of set size and experiment also reached significance, indicating a slightly different relation of the mean RTs at the medium set size to the mean RTs at the small and large set sizes in the two experiments, see *Figure 3* and *Figure 5*.) However, the interaction between congruent versus incongruent priming, the linear trend of set size and experiment did not reach significance.

Linear Mixed Model Analyses. RTs were analyzed in an LMM analysis as a function of priming condition (coded as +1 congruent; -1 incongruent), set size (i.e., 2 / 8 /16), and experiment (Experiment 2 was coded as +1; Experiment 3 was coded as -1), including all interaction terms. Across the two experiments, set size had a weight of b = 19.9 ms (SE = 1.0 ms), $t(22.92) = 20.38$, $p < .001$. Consistent with the results of the individual experiments, the interaction of set size and priming was significant, b = -1.3 ms (SE = 0.3 ms), $t(23.20) = -4.28$, $p < .001$. Furthermore, the interaction of set size and experiment was significant, b = -7.4 ms (SE = 1.0 ms), $t(22.92) = -7.53$, $p < .001$. Thus, when complex objects were presented, RTs increased by 14.8 ms more for each additional item in the display than when simple stimuli were presented. Consistent with the MANOVA results, the interaction between set size and priming was not further moderated by the experiment, b = 0.4 ms (SE = 0.3 ms), $t(23.20) = 1.50$, $p = .148$.

Discussion

In accordance with the results of Experiments 1 and 2, a significant congruency effect emerged in Experiment 3, and importantly, congruent priming resulted in flatter search slopes compared to incongruent priming, indicating more efficient search. As expected, the presentation of complex visual objects generally led to less efficient visual search compared to the use of simple color features. Moreover, we found a greater influence of primes with more complex material, as reflected in larger priming differences in Experiment 3 compared to Experiment 2, while the between-experiments comparison regarding the influence of primes on search efficiency did not reach significance.

Thus, Experiment 3 demonstrated that facilitation of search efficiency involved in cross-modal priming is not limited to simple information such as colors and their spoken denotations. Rather, the present results indicate that this mechanism is also applicable to search tasks in which the target is defined by several perceptual and semantic features, and the primes refer to semantic objects. There is evidence that attentional templates that guide selection can be established based on and influenced by object- and category-level semantic information (e.g., Yu et al., 2016). The present results indicate that object-level auditory information could be used as a source of guidance for visual attention. Overall, these findings are promising as they suggest that cross-modal influence on attentional guidance is effective on complex object representations, representing an important step toward testing this process with regard to real-world search behavior. Furthermore, Experiment 3 used the smallest target set that allows our manipulation of semantic congruency, thus, two possible targets. Together with the results of Experiment 2, Experiment 3 supports the claim that the working memory load of searching for four possible targets is not a necessary condition for auditory priming to influence search processes.

General Discussion

While several lines of evidence indicates that task-irrelevant sounds can improve performance in visual search tasks when they are semantically congruent (as opposed to incongruent) to the target (e.g., Iordanescu et al., 2008; Mahr & Wentura, 2014, 2018), the underlying processes remained elusive. In the present series of experiments, varying set size of the search displays allowed us to test the involvement of the following proposed processes: First, a performance advantage with congruent sounds might be attributed to a general mechanism proposed for semantic priming in settings with attended targets, namely facilitation of target encoding (e.g., McNamara, 2013), leading to improved target identification. Second, beyond facilitation of post-search processes, we proposed that cross-modal priming might be considered a source of guidance for selective visual attention (i.e., in terms of the guided search framework; Wolfe, 1994, 2021). Thus, we hypothesized that auditory primes would enhance search efficiency via guidance of attention. Thus, we proposed that cross-modal priming could be considered a source of guidance for selective visual attention (i.e., in terms of the guided search framework; Wolfe, 1994, 2021). Taken together, the present study provides direct evidence that semantically congruent auditory primes increase visual search efficiency: Specifically, four experiments (including Experiment 1R; see *Appendix*) provided evidence for cross-modal enhancement of search efficiency by congruent priming, that is, flatter search slope for congruent compared to incongruent trials. This effect emerged despite the varying cognitive load of maintaining four (Experiments 1 and 1R) or two (Experiments 2 and 3) potential targets throughout the task. Furthermore, while

Experiments 1(R) and 2 used a simple feature to define targets, Experiment 3 showed that this effect generalizes to complex audio-visual object representations.

This pattern of findings is only explainable under the assumption of an additional process beyond the predominantly assumed target encoding facilitation explanation; namely, when the target location is uncertain, auditory primes act as a source of guidance for selective visual attention. Thus, the present findings are consistent with the notion that auditory semantic priming can alter the activation of the priority map in favor of the primed target (see, e.g., Wolfe, 2021), enhancing its visual salience. By which mechanism can cross-modal priming alter activation of the priority map and what role can multiple-target search play in this process? Tentatively, cross-modal semantic priming might increase activation of the priority map in favor of the primed target by affecting the representation of the guiding template, similar to what has been proposed for other sources of attentional guidance (e.g., for repetition priming; see, Wolfe, 2021). Thereby, the observed results can be related to recent (unimodal) visual search research for multiple targets (for a review, see Ort & Olivers, 2020). Dominantly this research is focused on search for a small set of targets – small enough to assume their storage in working memory (but see, e.g., Drew et al., 2017, for multiple-target search with large sets). Since there were only two to four possible target items in our experiments, it is likely that the relevant target items were actively maintained in working memory as well (Cowan, 2001).[11] As the target features remained constant during the task (Experiment 1) or during a task block (Experiments 2-3), long-term memory might also

[11] We acknowledge that on average the capacity for simple colors is estimated to be in the range of only three to four items by using the change-detection paradigm with changing colors from trial to trial (Luck & Vogel, 1997, 2013; Vogel & Awh, 2008). Nevertheless, given the present conditions (i.e., permanent usage of the same target items in an extended period) it seems plausible to assume storage of up to four items.

contribute to the representation of attentional templates. However, in case of multiple target search, there is evidence that the role of working memory remains present even when target features are held constant for an extended period (Grubert et al., 2016).

As a possible mechanism, the auditory primes in the present study might exert their influence on attentional guidance by (relatively) prioritizing a search template (i.e., guiding template in the terminology of Wolfe, 2021) corresponding to the primed item. Such an interpretation would be consistent with the two currently predominant frameworks on the role of working memory in attentional guidance. Specifically, there is a current debate in this field on whether multiple target representations held in working memory can guide attention simultaneously (e.g., Bahle et al., 2020; Beck et al., 2012; Kerzel & Witzel, 2019) or whether only a single item at a given point in time has this attention-guiding capacity (e.g., Olivers et al., 2011). The latter position refers to theories of working memory (Garavan, 1998; McElree, 2006; Nee & Jonides, 2011; Oberauer, 2002; Olivers et al., 2011) claiming that although working memory may hold active several items, only a single item can directly influence perception and the deployment of attention (i.e., the "focus of attention"; Oberauer, 2002). It has been shown that this prioritized item has the potential to guide visual attention by increasing visual sensitivity for the corresponding visual stimulus (Desimone & Duncan, 1995; Olivers et al., 2006). The prioritized working memory item can thus be regarded as an active template that directly resonates with corresponding visual input (Olivers et al., 2011). In contrast, the multiple-item-template hypothesis assumes that, at least under certain conditions (e.g., using simple features), more than a single attentional template can be established and guide search processes (e.g., Kerzel & Witzel, 2019). Importantly, theoretical and empirical work supporting this notion do allow for graded activity levels of the search templates guiding selection (Bahle et al. 2020; see also Kerzel & Witzel, 2019) and therefore for differences in search performance between them (e.g., to account for switch costs, that is, slower responses if the target in trial n is different from the target in trial n-1; Ort et al., 2017).

Applied to the present context, congruent auditory priming might bring the corresponding working memory item into the prioritized status (i.e., “focus of attention”), thereby allowing that item to guide visual attention (e.g. Olivers et al., 2011). Alternatively, from the perspective of a genuine multiple-target search with a potential asymmetry in the activation of multiple attentional templates (e.g., Bahle et al. 2020), the congruent prime might simply cause an increased activity level of the corresponding item (instead of the all-or-none change of status). For either rationale, we would expect search for the (relatively) prioritized item to be more efficient. Because our experiments were not designed to decide the debate between these two frameworks, we acknowledge the ambiguity of our results in this regard.

The results of the neutral priming condition compared to the congruent and incongruent conditions may give further hints on the underlying mechanisms. With exception of Experiment 1(R), we found an approximately balanced distribution of the congruency effect into costs and benefits (see *Figures 2*, *3*, and *5*). These results fit with a guidance of visual search by the primed and therefore prioritized target: It has beneficial effects in the congruent case in terms of search efficiency. However, it also has detrimental effects for those targets that do not correspond with the currently prioritized item, as we found costs associated with incongruent trials. It is up to further research to explore whether there is genuine inhibition of the non-prioritized working memory items or whether the detrimental effect is a by-product of the prioritization of the primed item (e.g., focusing on a distractor that has some similarity with the prioritized item). The former assumption certainly fits better to the single-item-template hypothesis (Olivers et al., 2011); the latter is compatible with both the single-item-template and the multiple-item-template hypotheses (Beck et al., 2012). As said, the balanced distribution of costs and benefits does not hold for the smaller set sizes of Experiment 1 (see *Figure 2*) and for the larger set sizes of Experiment 1R (see *Table A1*). We will return to this point below. However, the interpretation of the cost-benefit results should be treated with some caution, as the choice of neutral baseline condition may affect the interpretation of these difference scores (for a discussion regarding the response competition setting, see, e.g., MacLeod, 1991; see also Mahr & Wentura, 2014).

In all experiments, the search processes were inefficient, as indicated by the large influence of set size on search times. First, this is evidently related to the highly heterogeneous distractor items and, second, to the cost of keeping multiple attentional templates active. Consistent with this, Experiment 2 with a reduced target set was indeed associated with more efficient visual search than Experiment 1R (i.e., the small-*N* replication of Experiment 1 that can be directly compared with Experiment 2). Regarding the general influence of primes, priming differences were numerically larger with a larger target set of four than a smaller set of two potential targets; however, the comparison between experiments just failed to be significant. The difference between the search slopes for congruent versus incongruent priming was significantly larger when using a larger target set. Thus, with the larger target set, there was an increased influence of primes on search efficiency. This finding is consistent with two different theoretical assumptions. First, it may be related to a higher working memory load with a larger set of potential targets, which could lead to a reduced inhibition of the influence of task-irrelevant stimuli in general (e.g., Lavie, 2010). Indeed, this possibility was the reason for conducting Experiment 2: We wanted to test whether the influence of auditory primes on search efficiency holds even with a considerably lower working memory load. Now, that we see that the difference between Experiment 1R and Experiment 2 is only a gradual one with respect to the most important result (i.e., the interaction between congruent versus incongruent priming and the linear trend of set size), we should note that, second, this finding is also consistent with our general theoretical assumption that auditory priming could bring the corresponding working memory item into a (relatively) prioritized state. If we would simply assume that in a neutral context (i.e., in the present setting, in the neutral priming condition) it is at random which one of the possible targets is in a prioritized state, and that target-related primes would always bring the corresponding item into a prioritized state, then in the four-target setting (i.e., Experiment 1R), the congruent condition, associated with a 100% congruency between the prioritized item and the presented target, would stand in contrast to a 25% congruency in the neutral condition. However, in the two-target setting (i.e., Experiment 2), this would be 100% congruency in the congruent condition versus 50% congruency in the neutral condition. Thus, the benefit component of priming would be increased in the four-target condition. Thus, our theoretical assumption provides a plausible alternative, especially considering the numerical pattern (i.e., the benefit component was descriptively larger in Experiment 1R compared to Experiment 2 at the larger set sizes, see *Table A1* and *Table 2*). However, this assumption needs to be corroborated by future research.

Furthermore, multiple target search is suggested to involve a further process not yet discussed: An internal search process of target templates (i.e., veridical target representations in the activated long-term memory; Wolfe, 2021) that serves to decide whether the selected item matches a possible target (e.g., Cunningham & Wolfe, 2014; Wolfe, 2012). Several lines of evidence indicate that semantic and further contextual influences can affect the memory search process. These include the categorical status of items (e.g., Cunningham & Wolfe, 2014; Shang et al., 2024), evidence for the influence of contextual cues that signal which memory set is currently relevant (e.g., Boettcher et al., 2013; see Boettcher et al., 2018, for evidence restricted to the setting where the relevant set remains constant for several trials), or spatial and temporal associations that can help to facilitate the visual and memory search process (e.g., Wiegand et al., 2021, 2024). It is up to further research to determine if cross-modal semantic priming may also influence memory search by increasing the accessibility of the primed target template.[12] This process would lead to more efficient memory search on congruent trials, because the most active target template already matches the target. Note, however, that this process would not explain an increased influence of primes on search efficiency with a larger target set, as it would predict a benefit that is independent of set size: For non-match decisions for the selected distractors, it would make no systematic difference whether the target-congruent or an incongruent target template is in a prioritized state.

Similar to Experiment 2, Experiment 3 also used a target set of two, but it used complex visual objects as targets and object-level semantic information as primes. This experiment replicated our main finding: Priming effects increased significantly with set

[12] We thank a Iris Wiegand for providing us with this suggestion.

size, indicating increased search efficacy by congruent priming. Thus, this effect occurred even though the facilitation of guidance could not be based on a well-defined perceptual feature, but rather on object-level auditory information. As discussed above, auditory primes may influence attentional guidance by (relatively) prioritizing the search template that corresponds to the primed item, thereby enhancing its visual salience. Given the growing empirical evidence (e.g., Robbins & Hout, 2020; Yu, et al., 2016) and theoretical considerations (e.g., Yu et al., 2023) suggesting that search templates guiding selection can be derived from and shaped by complex semantic information, it seems reasonable to assume that both feature- and object-level semantic primes can exert their influence by facilitating sensory gain, either for a specific feature or for a coarse representation of the primed object's features.

As shown by the comparison of Experiments 2 and 3, the use of more complex stimuli led, as expected, to a generally less efficient visual search and also to a greater influence of primes. The latter was indicated by larger priming differences for object search than for color search, while the difference in facilitation of search efficiency did not reach significance. It is an intriguing question for future research to corroborate this exploratory analysis and to test whether the relative contribution of semantic priming to guidance and post-search processes (e.g., target encoding, memory search) differs between feature- and object-level semantic primes.

In addition, there are three other noteworthy aspects of our study: First, primes influenced search processes even when they were presented briefly (i.e., time-compressed) with short SOAs and had no predictive information about the subsequent target, supporting the notion that the influence of primes is "automatic" in nature (i.e., fast, non-strategic). Second, in three experiments, we successfully adapted a small-*N* approach in which a large number of observations were produced by a relatively small number of participants. This approach might be fruitful for further research investigating perceptual and attentional effects that are expected to exhibit low interindividual variance and are little influenced by practice (e.g., Miller, 2023). Third, while Experiments 1 and 2 demonstrated facilitation of visual search efficiency when using simple stimuli, by using perceptually similar automotive symbols as targets and their spoken denotations as primes, Experiment 3 extended these findings to more complex audio-visual object representations. The results of Experiment 3 indicate that visual search facilitation by cross-modal priming is also applicable to search tasks in which the attentional template contains several perceptual and semantic features. Furthermore, these results have potential practical relevance: They suggest that speech warnings may be particularly useful in helping drivers to detect a relevant object (e.g., a potential hazard) in highly complex visual environments, where their visual attention is not initially focused on that object. It is up to future research to investigate whether cross-modal semantic priming may add benefit to the direct spatial information that might be provided by such a warning system. The fact that facilitation of visual search occurred with short, compressed auditory primes that preceded visual displays with brief SOA, and furthermore that this effect occurred even though attentive listening to the primes would not benefit the participants (i.e., primes and targets were uncorrelated), supports the notion that auditory warnings have the potential to "automatically" influence the deployment of visual attention. Thus, these results could provide an important basis for testing the role of speech warnings in guidance of visual attention in more realistic settings, such as dynamic driving simulations. Of course, further research should determine whether the effect is restricted to a limited set of possible targets or whether it is generalizable to large potential target pools (i.e., as in a complex driving environment).

Constraints on Generality of the Findings

The participants in our study were students at Saarland University, Germany, and therefore human. As our theoretical considerations and hypotheses relate to general human cognitive processes, this gave us the opportunity to test and possibly falsify our claims. This statement should not be confused with a claim that our results generalize to other human subpopulations.

Regarding the complexity of the stimuli, while Experiment 3 extended the findings to automobile symbols and their spoken denotations, future research should test the search efficiency hypothesis for more complex stimuli and real-world search behaviors. On a related note, it remains to be explored whether cross-modal semantic priming during visual search entails comparable processes when spoken primes naming simple features, objects, categories or characteristic sounds are used as primes. In addition, using more levels of set size may further help to corroborate the underlying relationship between priming and set size. Furthermore, as discussed earlier, further research is necessary to investigate the underlying mechanisms by which auditory primes exert their attention-guiding effect in the context of multiple-target search.

Conclusions

To conclude, while previous studies using cross-modal semantic priming of visual targets in distractor arrays have suggested that congruent (as opposed to incongruent) primes can enhance search efficiency, our study provided a test of this assumed underlying process. We showed that cross-modal priming leads to more efficient search, as indicated by flatter search slopes for congruent compared to incongruent trials. This finding is consistent with a shift in attentional prioritization in favor of the primed item, contributing to attentional guidance. This effect appeared to be robust across varying stimulus complexity and different levels of cognitive load, suggesting that it represents a general mechanism of audio-visual information processing. While further research should elucidate the exact mechanism by which auditory primes can influence attentional guidance in the context of multiple-target search, the present study represents an important step towards revealing the underlying processes, also of potential relevance to real-world applications.

Compliance with Ethical Standards and Conflict of Interest Statement

Funding. The research reported in this article was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG) to Dirk Wentura (WE 2284/9).

Conflict of interest. All authors declare that they have no conflict of interest.

Ethical approval. All procedures performed in the present study was conducted in accordance with the declaration of Helsinki (2013), the guidelines of the German Research Foundation (DFG), and the guidelines of the Ethics Committee of the Faculty of Empirical Social Sciences of Saarland University.

Informed consent. Informed consent was obtained from all individual participants included in the study.

References

Alexander, R. G., Nahvi, R. J., & Zelinsky, G. J. (2019). Specifying the precision of guiding features for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 1248-1264. doi:10.1037/xhp0000668

Bahle, B., Thayer, D. D., Mordkoff, J. T., & Hollingworth, A. (2020). The architecture of working memory: Features from multiple remembered objects produce parallel, coactive guidance of attention in visual search. *Journal of Experimental Psychology: General, 149*, 967-983. doi:10.1037/xge0000694

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1-48. doi:10.18637/jss.v067.i01

Beck, V. M., Hollingworth, A., & Luck, S. J. (2012). Simultaneous control of attention by multiple working memory representations. *Psychological Science, 23*, 887-898. doi:10.1177/0956797612439068

Belke, E., Humphreys, G. W., Watson, D. G., Meyer, A. S., & Telling, A. L. (2008). Top-down effects of semantic knowledge in visual search are modulated by cognitive but not perceptual load. *Perception & Psychophysics*, *70*, 1444–1458. doi:10.3758/PP.70.8.1444

Boettcher, S. E., Drew, T., & Wolfe, J. M. (2013). Hybrid search in context: How to search for vegetables in the produce section and cereal in the cereal aisle. *Visual cognition*, *21*, 678-682. doi: 10.1080/13506285.2013.844959

Boettcher, S. E., Drew, T., & Wolfe, J. M. (2018). Lost in the supermarket: Quantifying the cost of partitioning memory sets in hybrid search. *Memory & Cognition*, *46*, 43-57. doi:10.3758/s13421-017-0744-x

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, *1*, 1-20. doi:10.5334/joc.10

Burnham, B. R. (2010). Cognitive load modulates attentional capture by color singletons during effortful visual search. *Acta Psychologica*, *135*, 50–58. doi:10.1016/j.actpsy.2010.05.003

Calvert, G. A., Spence, C., & Stein, B. E. (2004). *The handbook of multisensory processes*. Cambridge, MA: MIT Press.

Causse, M., Imbert, J. P., Giraudet, L., Jouffrais, C., & Tremblay, S. (2016). The role of cognitive and perceptual loads in inattentional deafness. *Frontiers in human neuroscience*, *10*, 344. doi:10.3389/fnhum.2016.00344

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

Chen, Y.-C., & Spence, C. (2013). The time-course of the cross-modal semantic modulation of visual picture processing by naturalistic sounds and spoken words. *Multisensory Research, 26*, 371-386. doi:10.1163/22134808-00002420

Chen, Y.-C., & Spence, C. (2018a). Audiovisual Semantic Interactions Between Linguistic and Nonlinguistic Stimuli: The Time-Courses and Categorical Specificity. *Journal of Experimental Psychology: Human Perception and Performance, 44*, 1488-1507. doi:10.1037/xhp0000545

Chen, Y.-C., & Spence, C. (2018b). Dissociating the time courses of the cross-modal semantic priming effects elicited by naturalistic sounds and spoken words.

Psychonomic Bulletin & Review, 25, 1138-1146. doi:10.3758/s13423-017-1324-6

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

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407-428.

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

Cunningham, C. A., & Wolfe, J. M. (2014). The role of object categories in hybrid visual and memory search. *Journal of Experimental Psychology: General*, *143*(4), 1585-1599. doi: 10.1037/a0036313

De Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, *7*, 460–467. doi:10.1016/j.tics.2003.08.014

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193-222.

Drew, T., Boettcher, S. E. P., & Wolfe, J. M. (2017). One visual search, many memory searches: An eye-tracking investigation of hybrid search. *Journal of Vision, 17*. doi:10.1167/17.11.5

Driver, J., & Spence, C. (1998). Crossmodal attention. *Current Opinion in Neurobiology, 8*, 245-253. doi:10.1016/S0959-4388(98)80147-5

Duncan, J. (1989). Boundary Conditions on Parallel Processing in Human Vision. *Perception*, *18*, 457–469. doi:10.1068/p180457

Duncan, J., & Humphreys, G.W. (1989).Visual search and stimulus similarity. *Psychological Review*, *96*, 433-458. doi:10.1037/0033-295X.96.3.433

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*, 143-149.

Fernandez-Duque, D., & Posner, M. I. (1997). Relating the mechanisms of orienting and alerting. *Neuropsychologia*, *35*(4), 477-486. doi:10.1016/S0028-3932(96)00103-0

Giard, M. H., & Peronnet, F. (1999). Auditory-Visual Integration during Multimodal Object Recognition in Humans: A Behavioral and Electrophysiological Study. *Journal of Cognitive Neuroscience*, *11*, 473–490. doi:10.1162/089892999563544

Garavan, H. (1998). Serial attention within working memory. *Memory & Cognition, 26*, 263-276. doi:10.3758/bf03201138

Grubert, A., Carlisle, N. B., & Eimer, M. (2016). The control of single-color and multiple-color visual search by attentional templates in working memory and in long-term memory. *Journal of Cognitive Neuroscience*, *28*, 1947–1963. doi:10.1162/jocn_a_01020

Grubert, A., & Eimer, M. (2018). The time course of target template activation processes during preparation for visual search. *Journal of Neuroscience*, *38*, 9527-9538. doi:10.1523/jneurosci.0409-18.2018

Holcomb, P. J., & Anderson, J. E. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes, 8*, 379-411. doi:10.1080/01690969308407583

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*, *96*, B23–B32. doi:10.1016/j.cognition.2004.10.003

Hutchison, K. A., Balota, D. A., Neely, J. H., Cortese, M. J., Cohen-Shikora, E. R., Tse, C.-S., Yap, M. J., Bengson, J. J., Niemeyer, D., & Buchanan, E. (2013). The Semantic Priming Project. *Behavior Research Methods*, *45*, 1099–1114. doi:10.3758/s13428-012-0304-z

Iordanescu, L., Grabowecky, M., Franconeri, S., Theeuwes, J., & Suzuki, S. (2010). Characteristic sounds make you look at target objects more quickly. *Attention, Perception, & Psychophysics, 72*, 1736-1741. doi:10.3758/APP.72.7.1736

Iordanescu, L., Grabowecky, M., & Suzuki, S. (2011). Object-based auditory facilitation of visual search for pictures and words with frequent and rare targets. *Acta Psychologica, 137*, 252-259. doi:10.1016/j.actpsy.2010.07.017

Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review, 15*, 548-554. doi:10.3758/PBR.15.3.548

Jarmasz, J., & Hollands, J. G. (2009). Confidence intervals in repeated-measures designs: The number of observations principle. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 63*, 124-138. doi:10.1037/a0014164

Kerzel, D., & Witzel, C. (2019). The allocation of resources in visual working memory and multiple attentional templates. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 645-658. doi:10.1037/xhp0000637

Klapetek, A., Ngo, M. K., & Spence, C. (2012). Does crossmodal correspondence modulate the facilitatory effect of auditory cues on visual search? *Attention, Perception, & Psychophysics, 74*, 1154-1167. doi:10.3758/s13414-012-0317-9

Knoeferle, K. M., Knoeferle, P., Velasco, C., & Spence, C. (2016). Multisensory brand search: How the meaning of sounds guides consumers' visual attention. *Journal of Experimental Psychology: Applied, 22*, 196-210. doi:10.1037/xap0000084

Kristjansson, A., & Johannesson, O. I. (2014). How priming in visual search affects response time distributions: Analyses with ex-Gaussian fits. *Attention, Perception, & Psychophysics*, *76*, 2199-2211. doi:10.3758/s13414-014-0735-y

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017) lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. doi:10.18637/jss.v082.i13

Kvasova, D., Garcia-Vernet, L., & Soto-Faraco, S. (2019). Characteristic Sounds Facilitate Object Search in Real-Life Scenes. *Frontiers in Psychology, 10*, 2511. doi:10.3389/fpsyg.2019.02511

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

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences, 9*, 75-82.

Lavie, N. (2010). Attention, Distraction, and Cognitive Control Under Load. *Current Directions in Psychological Science*, *19*, 143–148. doi:10.1177/0963721410370295

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

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203. doi:10.1037/0033-2909.109.2.163

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

Mahr, A., & Wentura, D. (2018). Does a “stoplight!” improve processing a stoplight? Cross-modal influences of time-compressed spoken denotations on automotive icon classification. *Journal of Experimental Psychology: Applied, 24*, 92-107. doi:10.1037/xap0000158

Mastroberardino, S., Santangelo, V., & Macaluso, E. (2015). Crossmodal semantic congruence can affect visuo-spatial processing and activity of the fronto-parietal attention networks. *Frontiers in Integrative Neuroscience, 9*, 14. doi:10.3389/fnint.2015.00045

Maxfield, J. T., Stalder, W. D., & Zelinsky, G. J. (2014). Effects of target typicality on categorical search. *Journal of Vision*, *14*, 1. doi: 10.1167/14.12.1

McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *Psychology of Learning and Motivation: Advances in Research and Theory, Vol 46* (Vol. 46, pp. 155-200). San Diego: Elsevier Academic Press Inc.

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

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

Miller, J. (2023). How many participants? How many trials? Maximizing the power of reaction time studies. *Behavior Research Methods*. doi:10.3758/s13428-023-02155-9

Moors, A. (2016). Automaticity: Componential, causal, and mechanistic explanations. *Annual review of psychology*, *67*, 263-287. doi:10.1146/annurev-psych-122414-033550

Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory—Where are we now, and where should we go next? *Psychonomic Bulletin & Review*, *23*, 1316–1340. doi:10.3758/s13423-015-0982-5

Nagy, A. L., & Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, *7*, 1209–1217. doi:10.1364/JOSAA.7.001209

Nee, D. E., & Jonides, J. (2011). Dissociable contributions of prefrontal cortex and the hippocampus to short-term memory: Evidence for a 3-state model of memory. *Neuroimage, 54*, 1540-1548. doi:10.1016/j.neuroimage.2010.09.002

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.

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.

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences, 15*, 327-334.

Ort, E., Fahrenfort, J. J., & Olivers, C. N. L. (2017). Lack of Free Choice Reveals the Cost of Having to Search for More Than One Object. *Psychological Science, 28*, 1137-1147. doi:10.1177/0956797617705667

Ort, E., & Olivers, C. N. L. (2020). The capacity of multiple-target search. *Visual Cognition*, 1-26. doi:10.1080/13506285.2020.1772430

Reali, F., Spivey, M. J., Tyler, M. J., & Terranova, J. (2006). Inefficient conjunction search made efficient by concurrent spoken delivery of target identity. *Perception & Psychophysics*, *68*, 959–974. doi:10.3758/BF03193358

Robbins, A., & Hout, M. C. (2020). Scene priming provides clues about target appearance that improve attentional guidance during categorical search. *Journal of Experimental Psychology: Human Perception and Performance*, *46*, 220-230. doi:10.1037/xhp0000707

Rouder, J. N., & Haaf, J. M. (2018). Power, dominance, and constraint: A note on the appeal of different design traditions. *Advances in Methods and Practices in Psychological Science*, *1*, 19-26. doi:10.1177/2515245917745058

Salverda, A. P., & Altmann, G. (2011). Attentional capture of objects referred to by spoken language. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1122–1133. doi:10.1037/a0023101

Shang, L., Yeh, L. C., Zhao, Y., Wiegand, I., & Peelen, M. V. (2024). Category-based attention facilitates memory search. *Eneuro*, *11*(2). doi: 10.1523/ENEURO.0012-24.2024

Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design. *Psychonomic bulletin & review*, *25*, 2083-2101. doi:10.3758/s13423-018-1451-8

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics, 73*, 971-995. doi:10.3758/s13414-010-0073-7

Spence, C., Senkowski, D., & Röder, B. (2009). Crossmodal processing. *Experimental Brain Research, 198*, 107-111. doi:10.1007/s00221-009-1973-4

Spivey, M. J., Tyler, M. J., Eberhard, K. M., & Tanenhaus, M. K. (2001). Linguistically Mediated Visual Search. *Psychological Science*, *12*, 282–286. doi:10.1111/1467-9280.00352

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Tabossi, P. (1996). Cross-moddal semantic priming. *Language and Cognitive Processes, 11*, 569-576. doi:10.1080/016909696386953

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*, 400–410. doi:10.1016/j.tics.2010.06.008

Tellinghuisen, D. J., Cohen, A. J., & Cooper, N. J. (2016). Now hear this: Inattentional deafness depends on task relatedness. *Attention Perception & Psychophysics, 78*, 2527-2546. doi:10.3758/s13414-016-1169-5

Tellinghuisen, D. J., & Nowak, E. J. (2003). The inability to ignore auditory distractors as a function of visual task perceptual load. *Perception & Psychophysics, 65*, 817-828.

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

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology-Human Perception and Performance, 34*, 1053-1065. doi:10.1037/0096-1523.34.5.1053

Van Moorselaar, D., & Theeuwes, J. (2023). Statistical learning within objects. *Psychological Science*, *34*, 501–511. doi:10.1177/09567976231154804

Vogel, E. K., & Awh, E. (2008). How to exploit diversity for scientific gain: Using individual differences to constrain cognitive theory. *Current Directions in Psychological Science, 17*, 171-176. doi:10.1111/j.1467-8721.2008.00569.x

Wiegand, I., Westenberg, E., & Wolfe, J. M. (2021). Order, please! Explicit sequence learning in hybrid search in younger and older age. *Memory & Cognition*, *49*(6), 1220-1235. doi: 10.3758/s13421-021-01157-2

Wiegand, I., Wolfe, J. M., Maes, J. H., & Kessels, R. P. (2024). Incidental learning of temporal and spatial associations in hybrid search. *Visual Cognition*, 1-16. doi: 10.1080/13506285.2024.2346991

Wolfe, J. M. (1994). Guided search 2.0 A revised model of visual search. *Psychonomic Bulletin and Review*, *1*, 202–238. doi:10 .3758/BF03200774

Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 14-73). Hove, UK: Psychological Press.

Wolfe, J. M. (2012). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science*, 23, 698-703. doi:10.1177/0956797612443

Wolfe, J. M. (2014). Approaches to Visual Search: Feature Integration Theory and Guided Search. In A. C. Nobre & S. Kastner (Eds.), *Oxford Handbook of Attention* (pp. 11-55). Oxford, UK: Oxford Univ. Press.

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*, 1060-1092. doi:10.3758/s13423-020-01859-9

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*, 0058. doi:10.1038/s41562-017-0058

Yu, C. P., Maxfield, J. T., & Zelinsky, G. J. (2016). Searching for category-consistent features: A computational approach to understanding visual category representation. *Psychological science*, *27*, 870-884. doi:10.1177/0956797616640237

Yu, X., Zhou, Z., Becker, S. I., Boettcher, S. E., & Geng, J. J. (2023). Good-enough attentional guidance. *Trends in Cognitive Sciences*, *27*, 391-403. doi:10.1016/j.tics.2023.01.007

Zhang, Q., & Li, S. (2020). The roles of spatial frequency in category-level visual search of real-world scenes. *PsyCh journal*, *9*, 44-55. doi:10.1002/pchj.294

Zweig, L. J., Suzuki, S., & Grabowecky, M. (2015). Learned face–voice pairings facilitate visual search. *Psychonomic Bulletin & Review, 22*, 429-436. doi:10.3758/s13423-014-0685-3

Tables

Table 1. Mean RTs (in ms) and error rates (in %) as a function of set size and semantic congruency conditions in Experiment 1; priming effects (PE), costs, and benefits (*SE* in parentheses)

	Semantic congruency condition								
Set Size	Neutral	Congr.	Incongr.	PE		Benefit		Cost	
RT									
2	734	709	727	18	(10)	25	(12)	-6	(11)
8	918	863	923	60	(15)	55	(15)	5	(13)
16	1111	1036	1191	155	(30)	75	(27)	80	(16)
Error rates									
2	7.8	7.6	8.6	1.0	(1.2)	0.1	(1.6)	0.8	(1.2)
8	6.8	7.2	8.6	1.4	(1.5)	-0.4	(1.4)	1.8	(1.0)
16	8.3	6.4	9.4	3.0	(1.3)	1.9	(1.3)	1.1	(1.1)

Note. PE = $RT_{incongruent} - RT_{congruent}$; Benefit = $RT_{neutral} - RT_{congruent}$; Cost = $RT_{incongruent} - RT_{neutral}$

Table 2. Mean RTs (in ms) and error rates (in %) as a function of set size and semantic congruency conditions in Experiment 2; priming effects (PE), costs, and benefits (*SE* in parentheses)

	Semantic congruency condition								
Set Size	Neutral	Congr.	Incongr.	PE		Benefit		Cost	
RT									
2	526	522	529	7	(2)	4	(1)	3	(2)
8	618	612	625	13	(3)	6	(3)	7	(3)
16	703	688	718	30	(10)	15	(7)	15	(5)
Error rates									
2	3.1	2.7	3.1	0.4	(0.5)	0.4	(0.3)	0.0	(0.3)
8	3.4	3.4	3.2	-0.2	(0.5)	0.0	(0.4)	-0.2	(0.3)
16	3.4	3.1	2.8	-0.3	(0.3)	0.3	(0.3)	-0.6	(0.4)

Note. PE = $RT_{incongruent} – RT_{congruent}$; Benefit = $RT_{neutral} – RT_{congruent}$; Cost = $RT_{incongruent} – RT_{neutral}$

Table 3. Results of the 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) × 2 (experiment: Experiment 1R vs. Experiment 2) MANOVA for repeated measures on RTs with prime-target relation and set size as within-participants factors and experiment as between-participants factor. We report the respective a priori planned contrasts below each main effect and interaction. *Effects across Exp. 1R and 2* refers to the respective comparisons without the experiment factor, while *Comparison between Exp. 1R and 2* refers to the respective comparisons including experiment factor.

		Effects across Exp. 1R and 2				**Comparison between Exp. 1R and 2**			
		F	df	*p*	η_p^2	*F*	df	*p*	η_p^2
prime-target relation		16.02	2,16	< .001	.667	2.34	2,16	.128	.227
congruent vs. incongruent		33.81	1,17	< .001	.665	4.34	1,17	.053	.203
neutral vs. congruent & incongruent		1.56	1,17	.229	.084	1.89	1,17	.187	.100
set size		160.47	2,16	< .001	.953	5.17	2,16	.019	.392
linear trend		249.01	1,17	< .001	.936	7.44	1,17	.014	.304
quadratic trend		2.90	1,17	.107	.146	0.25	1,17	.626	.014
prime-target relation × set size		6.28	4,14	.004	.642	1.63	4,14	.223	.317
congruent vs. incongruent	× linear trend	27.86	1,17	< .001	.621	5.27	1,17	.035	.237
	× quadratic trend	6.73	1,17	.019	.283	0.75	1,17	.400	.042
neutral vs. congruent & incongruent	× linear trend	3.10	1,17	.096	.154	3.37	1,17	.084	.166
	× quadratic trend	0.15	1,17	.701	.009	0.02	1,17	.898	.001

Table 4. Mean RTs (in ms) and error rates (in %) as a function of set size and semantic congruency conditions in Experiment 3; priming effects (PE), costs, and benefits (*SE* in parentheses).

Set Size	Semantic congruency condition								
	Neutral	Congr.	Incongr.	PE		Benefit		Cost	
RT									
2	533	519	543	24	(3)	14	(2)	10	(2)
8	711	690	734	44	(7)	21	(5)	24	(4)
16	913	878	949	71	(14)	34	(7)	36	(13)
Error rates									
2	2.8	2.9	3.2	0.3	(0.4)	-0.1	(0.4)	0.4	(0.3)
8	2.1	2.6	2.2	-0.4	(0.3)	-0.5	(0.3)	0.1	(0.3)
16	2.0	2.1	1.5	-0.6	(0.4)	-0.1	(0.3)	-0.5	(0.3)

Note. PE = $RT_{incongruent} - RT_{congruent}$; Benefit = $RT_{neutral} - RT_{congruent}$; Cost = $RT_{incongruent} - RT_{neutral}$

Table 5. Results of the 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) × 2 (experiment: Experiment 2 vs. Experiment 3) MANOVA for repeated measures on RTs with prime-target relation and set size as within-participants factors and experiment as between-participants factor. We report the respective a priori planned contrasts below each main effect and interaction. *Effects across Exp. 2 and 3* refers to the respective comparisons without the experiment factor, while *Comparison between Exp. 2 and 3* refers to the respective comparisons including experiment factor.

		Effects across Exp. 2 and 3				**Comparison between Exp. 2 and 3**			
		F	df	p	η_p^2	F	df	p	η_p^2
prime-target relation		33.33	2,22	< .001	.752	7.67	2,22	.003	.411
congruent vs. incongruent		63.65	1,23	< .001	.735	14.34	1,23	< .001	.384
neutral vs. congruent & incongruent		0.01	1,23	.938	.000	0.01	1,23	.930	.000
set size		244.88	2,22	< .001	.957	28.60	2,22	< .001	.722
linear trend		440.60	1,23	< .001	.950	59.02	1,23	< .001	.720
quadratic trend		1.32	1,23	.262	.054	5.90	1,23	.023	.204
prime-target relation × set size		9.37	4,20	< .001	.652	2.11	4,20	.118	.297
congruent vs. incongruent	× linear trend	21.14	1,23	< .001	.479	2.55	1,23	.124	.100
	× quadratic trend	1.11	1,23	.302	.046	0.08	1,23	.781	.003
neutral vs. congruent & incongruent	× linear trend	0.25	1,23	.619	.011	0.17	1,23	.687	.007
	× quadratic trend	0.36	1,23	.553	.016	0.07	1,23	.796	.003

Figures

Figure 1. Trial sequence of Experiments 1 and 2 (not drawn to scale; example: incongruent trial; set size 8). The prime word (here: “red”) was presented via headphones 90 ms prior to the visual target (here: yellow). In Experiment 1, participants had to search for four potential target colors in each block. In Experiment 2, the set of relevant target colors was reduced to two colors in each block. The possible target and non-target colors are depicted in *Figure A1.* The correct response in this example is “up” (because the gap is in the upper part of the yellow circle).

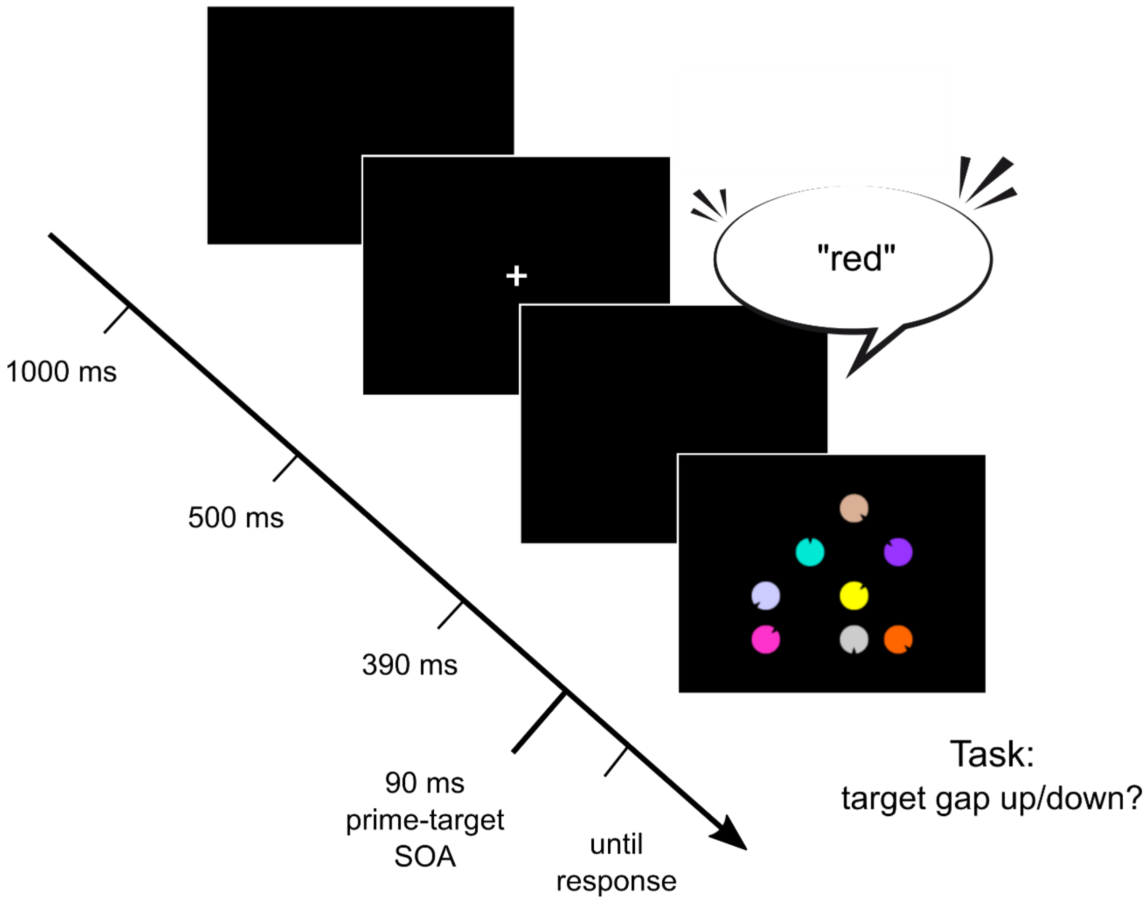


Figure 2. Mean RTs as a function of set size and priming condition in Experiment 1. Error bars are 95% within-subject confidence intervals (Jarmasz & Hollands, 2009) for the 2 (priming condition: congruent vs. incongruent) × 3 (set size) interaction effect.

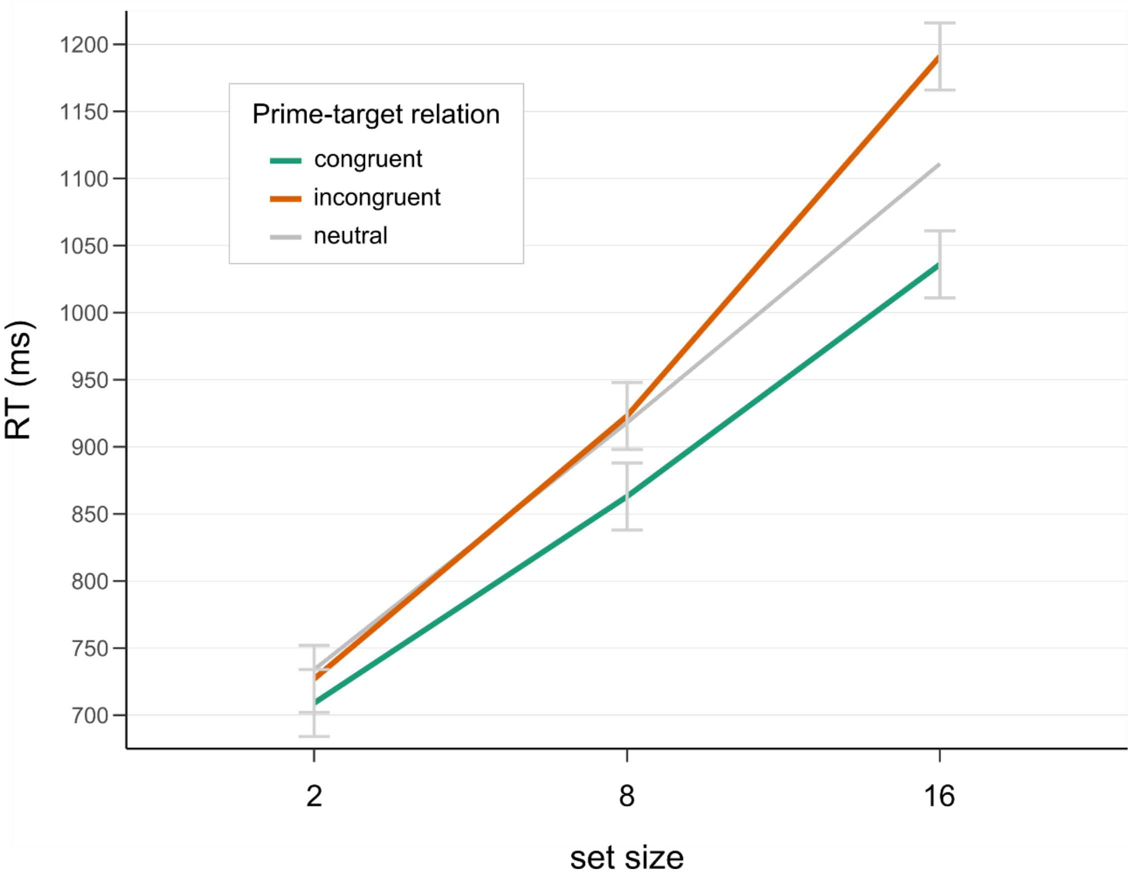


Figure 3. Mean RTs as a function of set size and type of priming in Experiment 2. Error bars are 95% within-subject confidence intervals (Jarmasz & Hollands, 2009) for the 2 (priming condition: congruent vs. incongruent) × 3 (set size) interaction effect.

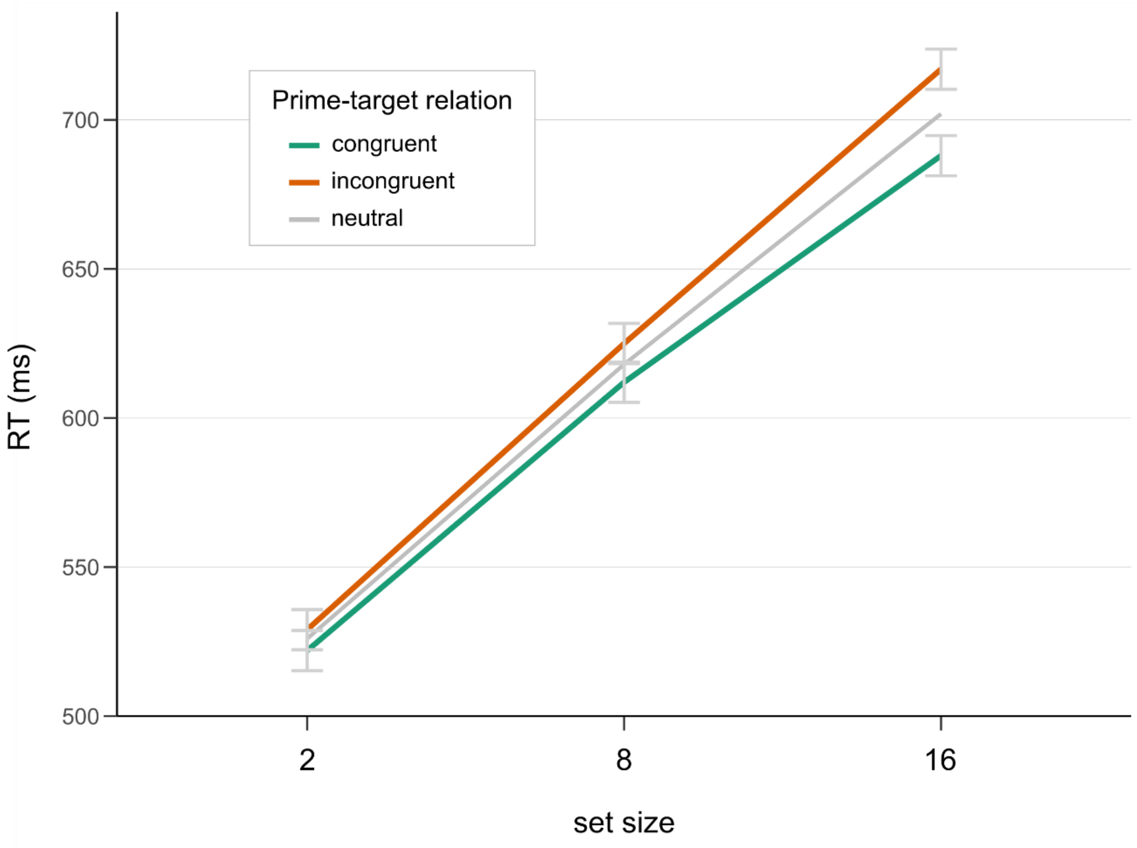


Figure 4. Trial sequence of Experiment 3 (not drawn to scale; example: congruent trial; set size 8). The prime word (here: “ambulance”) was presented via headphones 100 ms prior to the visual target. In each block, two symbols from the possible four target stimuli were presented as potential targets. The correct response in this example is “left” (i.e., the symbol of the ambulance is facing to the left).

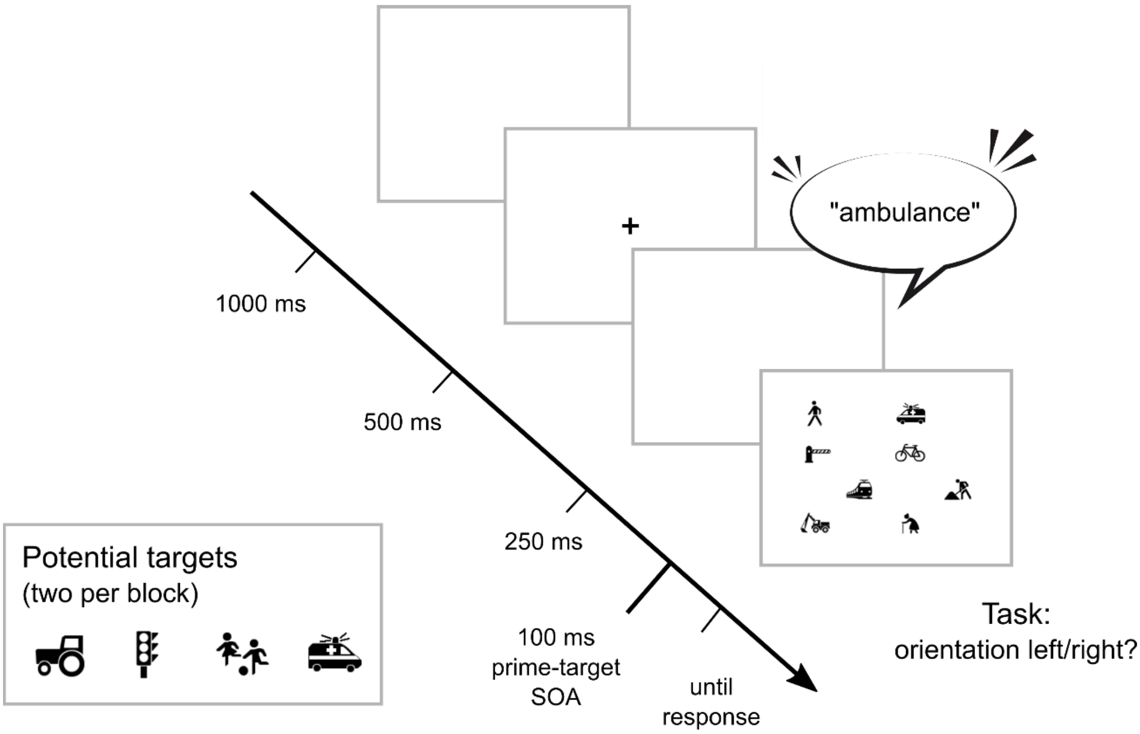
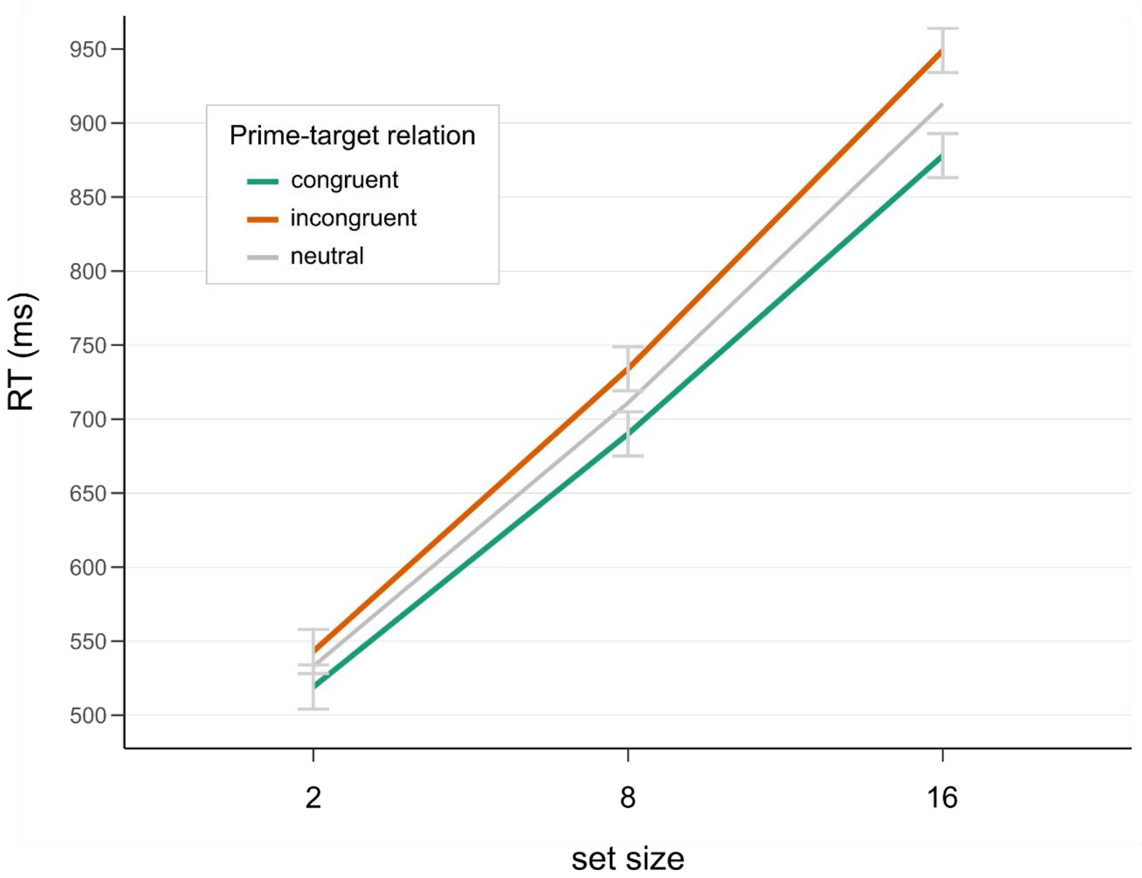


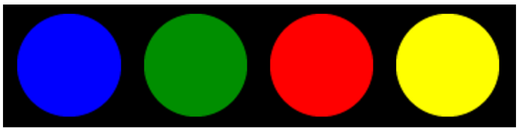
Figure 5. Mean RTs as a function of set size and type of priming in Experiment 3. Error bars are 95% within-subject confidence intervals (Jarmasz & Hollands, 2009) for the 2 (priming condition: congruent vs. incongruent) × 3 (set size) interaction effect.



Appendix

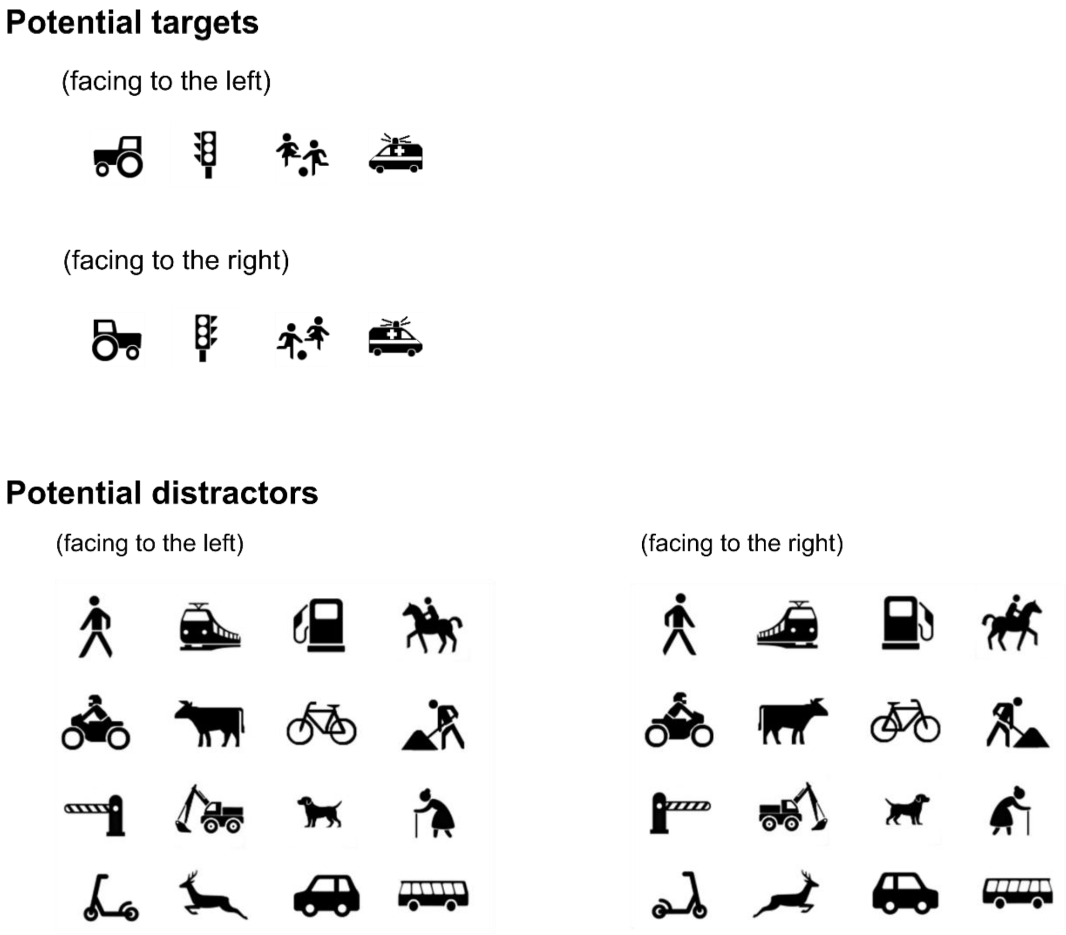
Figure A1. Visual stimuli used in Experiments 1 and 2. In Experiment 1, participants had to search for the four potential target colors in each block. In Experiment 2, the set of relevant target colors was reduced to two colors. In each block, the two relevant target colors were selected from the set of the four depicted target colors.

Target stimuli



Distractor stimuli

Figure A2. Visual stimuli used in Experiment 3. In each block, two symbols from the possible four target stimuli were presented as potential targets. Sixteen further symbols served as distractor stimuli. For each target and distractor symbol, there was a version that faced to the left, and one that faced to the right.



Experiment 1R – Replication of Experiment 1 using a small-*N* approach

Because the data collection of Experiment 2 was conducted during the COVID-19 pandemic, recruiting a sufficient sample size was not realistic using a typical testing approach (i.e., in the present case: recruitment of $N = 45$ and presenting 432 trials for each participant; see power calculation of Experiment 2). Since power is derived from both the number of participants and the number of repeated measurements per participant and condition, under certain assumptions, the number of observations can be distributed among participants and trials without compromising power (e.g., Rouder & Haaf, 2018; Brysbaert & Stevens, 2018). Of course, the following assumptions must be met for adopting a small-*N* approach, in which a large number of observations is produced by a relatively small number of participants: (1) The effect is stable over a large sequence of trials; (2) the effect is rather homogeneous across participants (Brysbaert & Stevens, 2018). Furthermore, Rouder and Haaf (2018) suggested that decreasing the sample size can be plausibly compensated by increasing the number of trials with little or no loss of power in the case of most RT-based cognitive effects, as also demonstrated by their simulations (see also the simulation studies of Brysbaert & Stevens, 2018, supporting this notion).While this discussion is gaining increasing attention, there is also growing evidence that paints a broader picture about the feasibility and limitations of this approach. By selecting subsets of participants and trials from large real datasets, Miller (2023) has demonstrated that a simple trade-off between the number of participants and the number of trials per participant has its limitations regarding many RT-based effects, for which the effect size typically decreases and the standard error of the effect increases with practice (e.g., word/non-word effect in lexical decision tasks; e.g., Hutchiton et al. 2013). Importantly, with regard to the semantic priming effect, the study of Miller (2023) found no evidence of such a limitation when the reduced sample size was compensated by an increased number of trials. To test the feasibility of this approach with respect to semantic priming of visual search processes, we wanted to replicate the results of Experiment 1 using a small-*N* design. Therefore, the present experiment was intended to be an exact replication of Experiment 1, except for adapting this strategy.

Method

Participants

Participants were 6 students (5 women, 1 men) from Saarland University, who took part in the experiment in exchange for 50 Euro. The median age was 23 years (ranged 21 to 26). All were native speakers of German and had self-reported normal hearing and normal or corrected-to-normal vision.

Experiment 1 had 240 trials per participant and found an effect size of $d_Z = 0.65$ for the most important effect, the interaction between the congruent versus incongruent priming conditions and the linear trend of set size. To detect an effect of $d_Z = 0.65$ with $\beta = .95$ ($\alpha = .05$, one-tailed) a sample size of $N = 28$ is needed. Thus, replicating the effect with $N = 28$ would yield a total number of observations of 6720. In the present study, $N = 6$ participants were recruited and they worked through a total number of 1920 trials each, resulting a total number of 11520 observations for the sample.

Design, Materials, Procedure

Materials, experimental procedure, design, data preparation, and exclusion criteria were identical to that of Experiment 1; except that each participant participated in eight runs of the experimental procedure of Experiment 1 (i.e., 240 trials each). Specifically, each participant was tested in two testing sessions that took place on two consecutive days or on two days with a day apart. Each testing session consisted of four runs of the procedure of Experiment 1.

Results

RTs of correct responses (mean error rate was 4.2%) were analyzed after outliers were discarded using the same criteria as in Experiment 1; this led to the exclusion of 4.9% of trials. Mean RTs and ERs as well as the difference scores of priming, benefits and costs for all set sizes are reported in *Table A1*.

Repeated Measures MANOVA

In line with the analysis of Experiment 1, we analyzed RTs in a 3 (prime-target relation: neutral vs. congruent vs. incongruent) × 3 (set size: 2 vs. 8 vs. 16) MANOVA for repeated measures (e.g., Dien & Santuzzi, 2005; O'Brien & Kaiser, 1985). We a priori focused on the contrast between congruent and incongruent conditions regarding the prime-target relation (see above; see also Mahr & Wentura, 2014, 2018), and, as we expected a linear relationship between set size and search RTs, the linear trend is a priori defined as the effect of main interest regarding set size.

The overall main effect of priming was significant, $F(2,4) = 8.24$, $p = .038$, $\eta_p^2 = .805$. The contrast of congruent vs. incongruent trials was also significant, $F(1,5) = 16.72$, $p = .009$, $\eta_p^2 = .770$, as expected, with faster responses in congruent trials than in incongruent trials (see *Table A1*). The main effect of set size was significant, $F(2,4) = 35.42$, $p = .003$, $\eta_p^2 = .947$. Search RTs increased with increasing display size, $F(1,5) = 58.80$, $p < .001$, $\eta_p^2 = .922$. (As expected, the quadratic trend was not significant, $F < 1$.) The interaction between prime-target relation (i.e., including the neutral condition) and set size was not significant, $F(4,2) = 1.98$, $p = .363$, $\eta_p^2 = .798$. Importantly, replicating the facilitation of visual search by congruent priming, there was a significant interaction between the congruent versus incongruent priming and the linear trend of set size, $F(1,5) = 18.29$, $p = .008$, $\eta_p^2 = .785$. (Other interaction contrasts were not significant, $F[1,5] = 5.70$, $p = .063$, $\eta_p^2 = .533$, for the interaction of congruent vs. incongruent priming and the quadratic trend of set size, all other *F*s < 2.28, *p*s >.191.).

Descriptively, mean RTs were shorter in the present experiment compared to Experiment 1 in each set size condition; this may reflect practice effects on the task during the long test session (see *Table 1* and *Table A1.*). Accordingly, the mean error rate in the present experiment was also numerically lower (4.2%) than in Experiment 1 (8.3%). Consistent with the lower mean RTs, the mean RT-differences between incongruent and congruent conditions were also numerically smaller for all set sizes compared to Experiment 1 (see *Table 1* and *Table A1*). However, each congruency effect was significantly different from zero, except in the smallest set size, which missed the conventional significance level; $t(5) = 1.72$, $p = .073$, $d_Z = 0.70$ for set size = 2; $t(5) = 4.05$, $p = .005$, $d_Z = 1.66$ for set size = 8; $t(5) = 2.47$, $p = .028$, $d_Z = 1.01$ for set size = 16. Priming differences as a proportion of base RT (i.e., RTs for neutral priming for a given set size) were 1.8% (*SE* = 1.1%), 3.9% (*SE* = 0.8%), and 8.0% (*SE* = 1.2%) for set sizes 2, 8, and 16, respectively. (Thus, numerically somewhat reduced compared to Experiment 1: 2.2%, 6.2%, and 12.9% for set sizes 2, 8, and 16, respectively.) Importantly, these relativized priming scores showed a significant linear increase with set size, $F(1,5) = 19.50$, $p = .007$, $\eta_p^2 = .796$. (See *Table A1* also for the difference scores of benefits and costs; benefits were significant only for the largest set size, $t[5] = 3.05$, $p = .028$, $d_Z = 1.25$; while *t*s < 1.91, *p*s >.114, for the lower set sizes. Costs were not significant for any of the set size conditions, *t*s < 2.41, *p*s >.060.)

For error rates, the corresponding MANOVA for repeated measures revealed no significant effects, *F*s < 1.57, *p*s > .313. For the contrasts of interest, the congruency effect was not significant, $F(1,5) = 0.87$, $p = .395$, $\eta_p^2 = .148$, nor was the interaction between congruent vs. incongruent trials and the linear trend of set size, $F(1,5) = 0.00$, $p = .957$, $\eta_p^2 = .001$. Thus, there was no evidence for a speed-accuracy tradeoff regarding the RT-effects (see also *Table A1*).

Linear Mixed Model Analyses

We again used linear mixed model approach to test if search slopes are moderated by the semantic congruency of the primes. RTs were analyzed as a function of the priming condition (coded +1 congruent, -1 incongruent; i.e., the neutral condition was discarded from the main analyses), set size (i.e., 2 / 8 /16), and their interaction (lmerTest package, Kuznetsova, et al., 2017; Bates, et al., 2015). We again fitted the maximal model (e.g., Barr et al., 2013). In this random slopes model, set size had a weight of $b = 17.4$ ms ($SE = 2.3$ ms), $t(5.00) = 7.42$, $p < .001$, indicating that with each additional element, search RTs increased by 17 ms. Importantly, the interaction of set size and priming was significant as well, $b = -2.1$ ms ($SE = 0.6$ ms), $t(7.21) = 3.25$, $p = .013$, indicating that search slopes differ for the two priming conditions: Slopes were flatter for congruent ($b = 15.3$ ms) than for incongruent trials ($b = 19.5$ ms).[13]

Discussion

The present experiment was intended to replicate the results of Experiment 1, importantly, by using an alternative testing approach of compensating for a lower

---

[13] The analysis including the neutral trials (with two contrast codes for priming and thus two interaction terms) yielded essentially the same results: A significant interaction contrast of congruent vs. incongruent priming with set size, $b = -2.1$ ms ($SE = 0.7$ ms), $t(7.02) = 3.21$, $p = .015$; while the interaction contrast of neutral vs. congruent/incongruent priming with set size was non-significant, $b = 0.8$ ms ($SE = 0.8$ ms), $t(9.46) = 1.12$, $p = .289$.

number of participants by a substantially higher number of trials per participant (see Brysbaert & Stevens, 2018; Rouder & Haaf, 2018). To sum up, this experiment replicated the main results of Experiment 1; that is, it revealed a significant reduction of the search slopes with congruent compared with incongruent priming.

In the present experiment, participants showed descriptively a better visual search performance compared to Experiment 1: Both mean RTs and errors were numerically smaller than in Experiment 1. In addition, search slopes were numerically less steep in the present experiment (with each additional item, search RTs increased by 17 ms in the replication of Experiment 1, whereas by 28 ms in Experiment 1). This may be due to the fact that participants had the opportunity to practice the task over a large number of trials, and may also be related to the characteristics of the sample (i.e., a lab experiment with two sessions of three hours net time each might attract different persons than a short experiment of less than 30 minutes).

The descriptively more efficient search is consistent with the numerically smaller priming effects in the present experiment (also as a proportion of base RT). However, priming differences (also relative to the base RT) increased with set size in the present experiment (in contrast to Experiment 1, the effect size of the priming differences did not increase from set size 8 to 16), and importantly, search slopes were also significantly flatter for congruent than for incongruent trials. In general, the replication of Experiment 1 provided further evidence for the search facilitation hypothesis with a much smaller number of participants and increased number of trials, opening the way for testing our further research questions by using this approach.

Table A1. Mean RTs (in ms) and error rates (in %) as a function of set size and semantic congruency conditions in the replication of Experiment 1 (Experiment 1R) using a small-*N* approach; priming effects (PE), costs, and benefits (*SE* in parentheses)

Set Size	Semantic congruency condition: Neutral	Congr.	Incongr.	PE		Benefit		Cost	
RT									
2	558	556	566	10	(6)	2	(5)	7	(6)
8	696	677	705	28	(8)	19	(10)	9	(4)
16	821	772	840	68	(15)	49	(16)	19	(12)
Error rates									
2	4.4	4.4	4.8	0.4	(0.6)	0.0	(1.2)	0.4	(0.9)
8	4.2	4.1	4.5	0.4	(1.0)	0.1	(0.6)	0.3	(0.8)
16	3.5	3.4	3.7	0.3	(1.0)	0.1	(1.1)	0.2	(1.0)

Note. PE = $RT_{incongruent} – RT_{congruent}$; Benefit = $RT_{neutral} – RT_{congruent}$; Cost = $RT_{incongruent} – RT_{neutral}$

Table A2. Deviations from the preregistrations and their reasons.

Exp.	Preregistration	Deviation	Reason
1-2	Error trials and RT outlier trials (i.e., RTs that were 1.5 interquartile ranges above the third quartile or below the first quartile, respectively, with respect to the individual distribution; Tukey, 1977) will be discarded.	We specified that we would discard RTs that were 1.5 interquartile ranges above the third quartile or below the first quartile (Tukey, 1977). There were no such outliers at the lower end of the individual distributions. However, we removed RTs below 150 ms (i.e., preparatory responses), which we forgot to mention in the preregistration. Specifically, 0.2% of all trials were excluded in Exp. 1. We also checked for preparatory responses in Exp. 2, there were no RTs below 150 ms.	It is common practice to remove RTs that are too fast to be interpreted as valid responses in the task (i.e., preparatory responses). Inclusion of these trials had no meaningful impact on results.
1-2-3	Error trials and RT outlier trials (i.e., RTs that were 1.5 interquartile ranges above the third quartile or below the first quartile, respectively, with respect to the individual distribution; Tukey, 1977) will be discarded.	Due to the marked differences between set sizes in terms of mean RTs and dispersion of RTs (see Tables 1, 3 and 4), the preregistered procedure resulted in unevenly distributed exclusion of trials in the three set size conditions. Therefore, we corrected our exclusion criterion: We excluded RT outliers that were 1.5 interquartile ranges above the third quartile of the individual RT-distribution *within each set size condition*. In detail, the % of trials that were excluded using our corrected vs. preregistered [in brackets] exclusion criteria are: **Exp. 1:** • Set size 2: 1.5% [vs. 0.3%] of trials • Set size 8: 1.8% [vs. 1.5%] of trials • Set size 16: 1.7% [vs. 4.9%] of trials **Exp. 2:** • Set size 2: 3.8% [vs. 0.8%] of trials	As can be readily seen, our preregistered criterion for identifying RT outliers that did not consider set size was flawed and resulted in a disproportionate exclusion of trials with respect to this condition (e.g., in Exp. 3, exclusion of up to nearly one-fifth of the trials in the largest set size condition while only 3.1% of trials in set size 2). We corrected this error for all experiments during our data processing.

		• Set size 8: 4.7% [vs. 4.3%] of trials • Set size 16: 5.0% [vs. 11.0%] of trials **Exp. 3:** • Set size 2: 5.8% [vs. 3.1%] of trials • Set size 8: 7.3% [vs. 6.4%] of trials • Set size 16: 7.6% [vs. 19.7%] of trials Since we initially excluded RT outliers in Exp. 1 according to our preregistered criterion, the effect sizes we oriented ourselves for the sample size planning of Exp. 2 and 3 are consistent with the results of this initial analysis. As to be expected, after correcting for this mistake in the exclusion criterion, the corresponding effect sizes are numerically larger (i.e., in Exp.1, $d_Z = 0.56$ and $d_Z = 0.65$ for the interaction between congruency and the linear contrast of set size using the preregistered and the corrected outlier criterion, respectively). Since we decided to orient ourselves on a reduced effect size of $d_Z = 0.50$ (i.e., a medium effect according to Cohen, 1988), this numerical discrepancy did not affect the actual determination of sample size.	
2	To detect an effect of $d_Z = 0.5$ with power 1-beta = .95 (alpha = .05, one-tailed) a sample size of $N = 45$ is needed.	The total number of observations with $N = 45$ and one experimental procedure (i.e., 432 trials) for each participant would have resulted in 19440 observations. We decided for the sample size of $N = 12$ participants, and in order to compensate for the reduction in the number of participants we increased the number of trials for each participant (i.e., 3456 trials), which resulted in 41472 observations.	As data collection took place during the COVID-19 pandemic, recruitment of the number of participants that was set in the Preregistration was not feasible. Therefore, we adopted a small-*N* approach, in which we compensated for a reduction in the number of participants by increased number of trials per participant (see main text; see also e.g., Brysbaert & Stevens, 2018). To test the feasibility of this approach for the present

			research, we conducted a replication of Exp. 1 using this approach (see *Appendix*).
3	A multilevel linear model will be fitted to the data (level 1 = trials, level 2 = participants). […] To test for Hypothesis 1 [i.e., RTs increase with increasing set size], the prediction of RTs through the set size factor will be evaluated. A significant main effect of set size with a positive regression weight would confirm Hypothesis 1 (RT set size 2 < RT set size 8 < RT set size 16). To test for Hypothesis 2 [i.e., congruency effect], it will be evaluated if prime-target congruency is a significant predictor for RTs. A negative regression weight would be expected to confirm Hypothesis 2 (RTcongruent < RTincongruent). […] To test for Hypothesis 3 [i.e., the congruency effect stated in Hypothesis 2 increases with increasing set sizes], the interaction term of prime-target congruency and set size is of relevance. To confirm Hypothesis 3 a significant interaction with a negative regression weight is expected.	To be consistent with the report of the results in Experiments 1 and 2, the test of Hypothesis 1 (i.e., congruency effect) was performed using repeated measures MANOVA. Furthermore, additionally to the preregistered LMM analyses, we report tests of Hypotheses 2 and 3 using repeated measures MANOVA. Testing Hypothesis 1 using LMM approach yielded consistent results with the MANOVA analysis reported in the main text. Specifically, to test the congruency effect using LMM approach, only the main effects of prime-target congruency and set size were included. The factor set size was centered at the set size of 8, so that the main effect of congruency can be interpreted as priming effect at set size 8. The random slopes model had the better fit to the data compared to the random intercepts model, $\chi2(5) = 333.39$, $p < .001$. Corroborating the results of the analyses reported in the main text, there was a priming effect of 46 ms (with a display size of 8), $b = -23$ ms ($SE = 3.3$ ms), $t(11.04) = -7.10$, $p < .001$.	We reported the MANOVA analysis for Hypothesis 1 and additionally to the LMM analysis for Hypotheses 2 and 3 to increase comparability with the results of Experiments 1 and 2.