Journal of Experimental Psychology:
Human Perception and Performance

ISSN: 0096-1523

2023, Vol. 49, No. 3, 327–354
https://doi.org/10.1037/xhp0001094

Top-Down Enhanced Object Recognition in Blocking and Priming Paradigms

Margit Scheibel[1] and Peter Indefrey[1, 2, 3]
[1] Department of Linguistics, Heinrich-Heine-University Düsseldorf
[2] Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen
[3] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Previous studies have demonstrated that context manipulations by semantic blocking and category priming can, under particular design conditions, give rise to semantic facilitation effects. The interpretation of semantic facilitation effects is controversial in the word production literature; perceptual accounts propose that contextually facilitated object recognition may underlie facilitation effects. The present study tested this notion. We investigated the difficulty of object recognition in a semantic blocking and a category priming task. We presented all pictures in gradually de-blurring image sequences and measured the de-blurring level that first allowed for correct object naming as an indicator of the perceptual demands of object recognition. Based on object recognition models assuming a temporal progression from coarse- to fine-grained visual processing, we reasoned that the lower the required level of detail, the more efficient the recognition processes. The results demonstrate that categorically related contexts reduce the level of visual detail required for object naming compared to unrelated contexts, with this effect being most pronounced for shape-distinctive objects and in contexts providing explicit category cues. We propose a top-down explanation based on target predictability of the observed effects. Implications of the recognition effects based on target predictability for the interpretation of context effects observed in latencies are discussed.

Public Significance Statement

This study highlights the role of object recognition in picture-naming tasks. Picture-naming paradigms like semantic blocking and category priming have been designed to investigate the stage of lexical-semantic processing in word production. The present study demonstrates that blocked/primed picture-naming tasks, however, influence not only lexical-semantic processing in a significant way, but also the efficiency of recognition processes. In two experiments, it is demonstrated that object recognition is facilitated in categorically related contexts compared to unrelated contexts, with this facilitation being most pronounced in the case of intra-category visual dissimilarity and when contexts provide explicit and reliable category cues to the target. We propose a top-down explanation of the observed effects based on target predictability. Implications of the observed recognition effects for the interpretation of picture-naming latencies are discussed.

Keywords: word production, object recognition, predictability, semantic blocking, semantic priming

When people are asked to name a depicted object, different steps of information processing are required. Visual information has to be perceived, interpreted, and transformed into conceptual information. The corresponding lexical information has to be retrieved, transformed into phonological and phonetic codes, and finally articulated as a sequence of speech sounds (cf. psychological theories on object recognition for the initial steps, e.g., Lindsay & Norman, 1977; and psycholinguistic theories on word production for the latter steps, e.g., Levelt et al., 1999). Picture naming thus involves a complex cascade of mental processes. Humans can, nevertheless, name depicted objects under a wide variety of circumstances, relatively effortless and fast (within a few hundred milliseconds) on a high level of accuracy.

Margit Scheibel https://orcid.org/0000-0001-5198-3770

This study was supported by the German Research Foundation (DFG, Collaborative Research Centre 991). Preliminary findings of the experiments were presented and discussed at the following conferences and workshops: Architectures and Mechanisms for language processing (AMLaP) 2017, International Workshop on Language Production (IWLP) 2018, Annual Meeting of the European Society of Philosophy and Psychology (ESPP) 2019, International Max Planck Research School (IMPRS) Conference 2020, and internal workshops and colloquia of the CRC 991. A complete list of stimuli, ratings collected in Pretest 3 (semantic similarity) and Pretest 4 (visual similarity), as well as the data collected in Experiments 1 and 2 and the analysis codes are publicly available at OSF and can be accessed at https://osf.io/r5kyp/?view_only=d6c19885fcee447483205c53071af5d7. The stimulus pictures cannot be made publicly available due to license restrictions of the original pictures but will be provided upon request. The authors declare to have no known conflicts of interest. The study was not preregistered.

Correspondence concerning this article should be addressed to Margit Scheibel, Heinrich-Heine-Universität, Abteilung für Allgemeine Sprachwissenschaft, Universitätsstr. 1, 40225 Düsseldorf, Germany Email: scheibel@uni-duesseldorf.de

The speed of picture naming is of particular interest to word production research. Picture-naming latencies are usually seen as a window to the relative difficulty of lexical access: the shorter the latencies, the easier the retrieval of the lexical representation (lemma) that corresponds to the depicted concept. A closer view on the multiple steps of information processing involved in a picture-naming process, however, casts doubt on the general validity of this reasoning. Picture-naming latencies could, in principle, also be dominated by one of the other steps of information processing. The open questions are under which circumstances naming latencies are determined by non-lexical processes, which non-lexical process dominates naming speed instead, and why.

Some conditions of semantic priming and semantic blocking paradigms are discussed as possibly being able to give pre-lexical effects an upper hand over lexical effects in picture-naming speed. This concerns in particular semantic priming conditions using long negative SOAs and first-cycle conditions of cyclic semantic blocking paradigms. For these conditions, it has often been argued that semantic effects could be the main determinant of naming latencies (e.g., Damian & Als, 2005; Lupker, 1988; Navarrete et al., 2014; Wheeldon & Monsell, 1994). An alternative proposal suggests that perceptual recognition effects could affect and possibly dominate picture-naming latencies in these particular conditions (e.g., Abdel Rahman & Melinger, 2007; Huttenlocher & Kubicek, 1983; Lin et al., 2022; Sperber et al., 1979). The notion that the time to process visually and to recognize pictured information could account, entirely or partly, for latency effects observed in some conditions of semantic priming and blocking paradigms would conflict with the purpose for which semantic priming/blocking tasks have been originally designed and are principally used in word production research. Behavioral priming/blocking studies, however, cannot rule out that perceptual recognition effects may become inadvertently dominant in naming latencies. The standard experimental design of semantic priming or blocking studies offers neither the possibility to check context-specific recognition difficulties nor to prove that naming latencies are strictly dominated by lexical-semantic processes, as intended.

Perceptual accounts of semantic context effects in naming latencies are mostly theoretical considerations as of yet. Although the notion as such is about 40 years old, little has been done to test a central assumption of perceptual accounts, that is, context manipulations by category priming and blocking would alter the perceptual demands of object recognition. In this paper, we address this under-investigated issue. We examined the context-specific ease of object recognition in semantic blocking (Experiment 1) and semantic priming paradigms (Experiment 2) by assessing the context-specific amount of visual information that is minimally sufficient for correct identification and naming of an object, for example, a horse. Perceptual accounts attributing shorter naming latencies to easier object recognition would hypothesize that perceptual recognition thresholds of objects should be lower in related contexts (homogeneous blocks, categorically primed conditions) as compared to unrelated contexts (heterogeneous blocks, unprimed conditions). Implications of our results for perceptual accounts and for the interpretation of latency effects reported in prior semantic priming/blocking studies will be discussed. The line of reasoning will go as follows. The detected object recognition thresholds will serve as an index to the amount of computation and time required to achieve object recognition in related and unrelated contexts. We will argue that the lower the recognition thresholds are, the more efficiently recognition processes seem to work and the faster object recognition may be reached.

Semantically Primed and Blocked Picture Naming: The Method and Typical Findings

Semantic blocking and semantic priming are two of the standard techniques to investigate lexical-semantic retrieval in word production. Both paradigms manipulate the semantic context in which to-be-named pictures are presented. In the semantic blocking paradigm, target pictures are presented in blocks. Each block employs a set of four or five different objects from either the same semantic category (homogeneous blocks; e.g., horse, cow, donkey, goat) or different semantic categories (heterogeneous blocks, e.g., horse, toaster, lemon, bicycle). In the cyclic version of the blocking paradigm, the set of objects is presented repeatedly in varying orders within a block such that each block consists of several presentation cycles, each presenting all objects of a set once. In semantic priming paradigms, including word/picture-priming tasks and picture-word-interference (PWI) tasks with negative SOAs, the semantic context is established by a usually not-to-be-named prime stimulus presented before the target (in PWI paradigms, before and during the target presentation). In related conditions, the prime stimulus refers to the target's semantic category, typically being a category coordinate, for example, cow—horse. In unrelated conditions, either an unrelated object, for example, toaster—horse, or no prime precedes the target picture.

Semantic blocking and category priming paradigms reliably produce context effects in naming latencies. These context effects can emerge as either semantic facilitation, an advantage of semantically related contexts relative to unrelated ones, or semantic interference, a term used to refer to the opposite effect: a disadvantage of related contexts relative to unrelated ones. Whether semantic facilitation or semantic interference is obtained seems to depend on design features beyond the mere related versus unrelated contrast, for example, on time parameters.

Semantic interference effects, that is, relatively longer naming latencies in related contexts, have been shown to emerge in priming/PWI paradigms when the prime/distractor is presented (nearly) simultaneously with the target (SOAs from −150 to +150 ms; e.g., Alario et al., 2000; Damian & Bowers, 2003; Damian & Spalek, 2014; Glaser & Düngelhoff, 1984) and in cyclic semantic blocking paradigms from the second naming cycle onwards (within-block repetition cycles; e.g., Belke et al., 2005; Damian & Als, 2005; Lin et al., 2022; Navarrete et al., 2012, 2014). The size of semantic interference effects is rather unvarying across studies and paradigms, usually amounting to some 20–50 ms.

Semantic facilitation effects, that is, relatively shorter naming latencies in related contexts, have been found in priming/PWI studies when the prime/distractor appears long before the target (SOAs between −300 and −1,000 ms; cf. Bürki et al., 2020) and in cyclic semantic blocking paradigms in the first cycle of a block (for a review, see Belke, 2017). In contrast to semantic interference effects, the size of semantic facilitation effects varies relatively greatly across studies. To give just some examples: Semantic blocking studies reported both no context effects and semantic facilitations up to about 50 ms for the first within-block cycle (e.g., Abdel Rahman & Melinger, 2007, 2011; Belke et al., 2005; Crowther & Martin, 2014; Damian & Als, 2005; Janssen et al., 2015; Lin et al., 2022; Meinzer et al., 2016; Navarrete et al., 2012, 2014; Python et al., 2018a). Previous semantic priming/PWI studies reported a similarly heterogeneous pattern of results with respect to long negative SOA conditions.[1] Python et al. (2018b), Sperber et al. (1979), and Zhang et al. (2016) obtained semantic facilitation effects of about 10–15 ms with word primes of category coordinates; Glaser and Düngelhoff (1984) observed numerical, though not significant facilitation of even 36 ms. Studies using picture primes of category coordinates reported facilitation effects from about 20 ms (La Heij et al., 2003, Experiment 2; Lupker, 1988, Experiment 1), to 30–50 ms (Huttenlocher & Kubicek, 1983, Experiment 1; Sperber et al., 1979) to even 100 ms (Humphreys et al., 1988, Experiment 2). Bajo (1988) obtained semantic facilitation effects above 50 ms for both word and picture primes and even larger facilitation effects if participants were instructed to concentrate on the semantic relation between prime and target. Semantic facilitation effects were also found with category-level word primes, for example, animal—dog. Davies et al. (1981) found that superordinate word primes facilitated naming by about 20 ms on average. Sperber et al. (1982; the results of the 11th graders are reported as an example) observed facilitation effects of about 30 ms if the target belonged to a nonperceptual category (category members are visually dissimilar) and facilitations of even about 50 ms if the target belonged to a perceptual category (category members are visually similar).

The Interpretation of Semantic Context Effects

Although the existence of semantic interference and semantic facilitation in picture-naming latencies is well established and beyond dispute, the interpretation of both phenomena is still a matter of debate in the word production literature.

Semantic interference effects are often interpreted as reflecting a lexical conflict (lexical competition accounts; e.g., Abdel Rahman & Melinger, 2009; Belke et al., 2005; Costa et al., 2005; Levelt et al., 1999; Roelofs, 1992). When a semantically related lexical candidate is nearly simultaneously perceived with the target (related primes/distractors presented with short/zero SOAs) or related lexical candidates enjoy high levels of residual activation due to repeated (co-)activation (within-block repetition cycles of homogeneous blocks), the target has to be selected from among a set of strongly co-active non-target nodes. Lexical competition accounts propose that all these active nodes enter into a competition for selection, which slows down the retrieval time for the target, the more, the larger the number of competitors and the higher their strength of activation. As the target reinforces the activation of related lexical candidates due to automatic activation spreading in conceptual networks, competition effects are greater in related than in unrelated contexts (cf. Abdel Rahman & Melinger, 2009, 2019). Accounts assuming a noncompetitive nature of lexical retrieval suggest that non-lexical effects interfere with naming. For the blocking paradigm, it has been proposed that incremental weakening of semantic-to-lexical mappings for repeatedly co-activated, but not retrieved words may slow down the naming times in within-block repetition cycles (e.g., Navarrete et al., 2014; Oppenheim et al., 2010). For picture-word paradigms, a conflict in a post-lexical, pre-articulatory response buffer has been proposed, a buffer to which words are supposed to have privileged access but target-related words (related distractors) may take longer to be excluded as a possible response than unrelated words (response exclusion hypothesis, e.g., Mahon et al., 2007). Despite the differing interpretations of inhibitory context effects, all current accounts of semantic interference concur in one important point: They all propose a post-perceptual basis of interference, thus they all interpret this effect as inherent to word production.

The interpretation of semantic facilitation effects is a bit more controversial. Semantic facilitation effects observed in semantic blocking and priming paradigms have been alternatively attributed to either the perceptual, semantic, lexical, or post-lexical stage. The crucial point is that, contrary to semantic interference effects, there is no general agreement on whether the effect has its basis in the word production system at all or whether it may reflect a facilitatory effect in a prelinguistic system, the visual recognition system, with the latter implying that the effect could possibly be a kind of artifact of the experimental method used to elicit word production, rather than an issue of word production. As a perceptual origin of semantic facilitation effects in the critical priming/blocking conditions would have important implications given the crucial role that has been attributed to context effects in discussions of theories of lexical access, we put the perceptual account of semantic facilitation to the test in this study. Before outlining our experiments and the hypotheses being tested, we first give an overview of the competing accounts of semantic facilitation that have been proposed in prior work.

Post-Perceptual Explanations of Semantic Facilitation Effects

There are several accounts sharing the view that semantic facilitation in naming latencies would arise at some stage in the word production system. However, the accounts disagree regarding the specific process that might get effectively quickened in related contexts. It has been suggested that either the conceptual-semantic access to the target (e.g., Abdel Rahman & Melinger, 2009; Belke et al., 2017; see also Lupker, 1988), its lexical access (e.g., Belke, 2017; Keefe & Neely, 1990), its semantic and lexical access (e.g., Navarrete et al., 2014), or post-lexical processes (Python et al., 2018a, 2018b) may be facilitated when the picture to be named is preceded by semantically related context item(s) compared to unrelated one(s). Nevertheless, despite the disagreements regarding the locus of facilitation effects, the underlying mechanism responsible for the higher efficiency of the respective process is almost consistently assumed to be semantic priming. The priming mechanism is, however, differently implemented in the different accounts.

Accounts proposing a conceptual or lexical locus of facilitation effects usually conceive of semantic priming as a short-lived, automatic mechanism based on spreading activation in the mental lexicon (e.g., Abdel Rahman & Melinger, 2009, 2019; Belke, 2013; Damian & Als, 2005; Navarrete et al., 2014; Wheeldon & Monsell, 1994). Bidirectional links between related conceptual representations and between each concept and its lexical counterpart allow activation to flow automatically within conceptual networks (via a common category node/direct links between related nodes, e.g., Levelt et al., 1999, or via shared semantic features, e.g., Dell et al., 2007; Oppenheim et al., 2010; Vigliocco et al., 2004). The result of spreading activation is

[1] As mentioned earlier in this section, for zero or short positive/negative SOAs the reversal, a semantic interference effect, is typically found (for a meta-analysis on SOA as a factor influencing the polarity of context effects in PWI paradigms, see Bürki et al., 2020).

that during the course of context processing a cohort of inter-related concepts and their lexical counterparts becomes automatically co-activated, which includes the target in related contexts and facilitates its access at the conceptual and/or lexical level when the target is encountered as compared to unrelated contexts in which no automatic pre-activation of the respective target representation takes place. Hence, such accounts suggest that the critical time savings in related contexts takes place during the lexicalization process, and originate from a mechanism inherent to it.

An alternative view is that semantic priming takes place strategically in a top-down fashion in related contexts. Strategic priming accounts suggest that participants may anticipate the target's semantic category in related contexts and exploit this a priori semantic information in order to facilitate target processing at a particular level, being alternatively the conceptual level (e.g., Belke et al., 2017), the lexical level (e.g., Belke, 2017; Keefe & Neely, 1990), or some post-lexical level (Python et al., 2018a, 2018b). Belke et al. (2017; see also Belke, 2017) argued that participants may notice the shared semantic category of the items in categorically homogeneous blocks over the course of the first two trials of a block and might use this knowledge to predict the category of the upcoming items of the block. Alternatively, the prime stimulus may be used to generate an expectancy about a set of related items likely encountered next (e.g., Keefe & Neely, 1990). The strategic pre-activation of likely exemplars of the anticipated category takes place in order to bias and hence facilitate processing at the critical processing level when the target has to be processed. As unrelated contexts are unpredictive to the target's semantic category, they do not allow for any strategic pre-activation, leading to relatively longer response times in unrelated contexts. Accounts implementing semantic priming as a strategic top-down mechanism based on categorical predictions suggest that the relative time savings in related contexts take place in the word production system and rely on knowledge that is stored in the mental lexicon, but the mechanism as such is not inherent to any process of the word production system. The priming mechanism is thought to be "under the subject's strategic control" (Neely et al., 1989, p. 1003); its operation hinges on whether and when participants notice the predictive potential of related contexts and how they put the context information to use in order to accelerate the naming response.

The question why semantic priming should matter only in some semantically blocked/primed conditions but not all, namely only in first within-block cycles and priming conditions with long negative SOAs, has been differently answered for the two paradigms. For the blocking paradigm, it has been argued that parallel inhibitory blocking effects may outweigh facilitatory priming effects from the second cycle onwards. When the same small set of related items has to be named repeatedly in close succession, lexical competition effects or incremental weakening effects at the semantic-to-lexical interface get enhanced, which cancels out priming effects in latencies in within-block repetition cycles. An absence of semantic facilitation in picture-word paradigms using short/zero SOAs has often been attributed to the mutual semantic priming of target and word stimulus in related conditions, with facilitation effects in latencies being only obtained when the target benefits more from the prime/distractor's priming than vice versa. Longer SOAs that allow for an earlier and deeper processing of the prime/distractor word (cf. Alario et al., 2000) apparently help the target to benefit from the prime word. Verbal information has been shown to access the conceptual representational level at a rather slow rate, relative to pictorial information (cf. Smith & Magee, 1980; see also Roelofs, 1992), as a result of which it may take comparatively longer for the spread of activation to produce sufficient priming input for the target. Strategic accounts suggest that a certain amount of time (around 200–250 ms) would be needed to generate semantic-taxonomic predictions and bring them strategically into play (cf. Neely & Keefe, 1989).

The strongest evidence that supports accounts assuming a form of semantic priming to be the driving force of facilitatory context effects in the critical priming/blocking conditions comes from studies demonstrating stronger facilitation with semantically close contexts than with semantically distant contexts (e.g., Lin et al., 2022; Navarrete et al., 2012) and studies demonstrating that facilitation effects become smaller or disappear when semantically related contexts are not strictly semantically related (e.g., Damian & Als, 2005; Navarrete et al., 2012). Despite the seemingly broad acceptance of semantic priming explanations, some uncertainty remains. Evidence that the basis of semantic facilitation must be post-perceptual is lacking or vague at best. Most of the data that are used to support the notion come from overall picture-naming times, which can in principle be influenced by a number of different processes, including perceptual ones. This inherent ambiguity of overall picture-naming times makes post-perceptual explanations of semantic facilitation still conjectural in part.

Perceptual Explanations of Semantic Facilitation Effects

The notion that facilitatory context effects could possibly reflect perceptual recognition effects, rather than effects related to word production, has a rather long history in the category priming literature and has been introduced to the semantic blocking literature in 2007 by Abdel Rahman and Melinger (for semantic priming paradigms, see e.g., Huttenlocher & Kubicek, 1983; Sperber et al., 1979; for semantic blocking paradigms, see Abdel Rahman & Melinger, 2007; Aristei et al., 2011; Lin et al., 2022). At the heart of perceptual accounts is the assumption that time savings in related contexts go back to entry-level processes for pictures: When the target picture is preceded by items from the same semantic category, the visual processing of the pictorial information and/or the conceptual identification/categorization of the pictured information may be facilitated. The consequence will be easier and faster recognition of the depicted object in related contexts compared to unrelated ones.

Current evidence in support of the notion that semantic priming/blocking may facilitate visual processing or conceptual interpretation of visual information in the critical conditions (first within-block cycles, long negative SOA conditions) is scant—not least because perceptual recognition processes are often ignored in word production research, and object recognition research that investigates perceptual processes employs other techniques to investigate context effects.[2]

[2] Context effects on visual object recognition are in principle well documented in the perception literature (for a review, see Bar, 2004), but studies investigating context effects on recognition processes usually manipulate visual scene contexts, that is, they present the target objects either in their usual surroundings (e.g., a hairdryer in a bathroom) or in unusual surroundings (e.g., a hairdryer in a forest). Semantic priming/blocking contexts differ from such visual scene contexts in that they provide taxonomic information on the target (rather than environment information), and they provide this information separately and before the target (rather than together with the target). These differences make it unclear whether one can compare known context effects on object recognition to blocking/priming effects on object recognition.

In the semantic priming literature, mainly two types of findings have been used to support perceptual accounts. The first type of evidence comes from priming studies demonstrating modality effects: Semantic facilitation is only obtained when the target is a picture (Huttenlocher & Kubicek, 1983) or it is at least substantially larger when the target is a picture rather than a word (e.g., Sperber et al., 1979); and semantic facilitation seems to be larger with picture primes than with word primes (Carr et al., 1982; Sperber et al., 1979; see also the review in section Semantically Primed and Blocked Picture Naming: The Method and Typical Findings, but see Lupker, 1988, Experiment 3, obtaining no significant effect of prime modality). The observation that the pictorial format seems to matter for the strength of semantic facilitation has spurred the idea that visual priming may account for or contribute to shorter latencies in primed conditions. Objects of the same semantic category have at least some visual features in common (cf. Lupker, 1988; Snodgrass & McCullough, 1986; Sperber et al., 1979). This could enable the prime stimulus to visually prime some object features of related targets, which facilitates visual processing of the target stimulus.

The second finding that has been used to support perceptual accounts is an interaction of category priming effects with effects of stimulus quality detected by Sperber and colleagues (Sperber et al., 1979, Experiment 1). In an attempt to test whether category priming can affect the initial processing level at which perceptual variables operate, Sperber et al. manipulated two variables in their priming study, the semantic relatedness of primes and the visual clarity of targets. They contrasted clear and visually degraded picture targets (the method of degradation produced a form of blurring) and observed that categorically related picture primes partially compensated the relative recognition disadvantage for degraded targets (priming effects for degraded targets were twice as large—112 ms —as those for clear targets—51 ms), indicating that both variables can apparently interact, thus likely operate at the same level. The authors suggested that the obtained facilitation effects in picture-naming latencies "result from a lowering of the evidence requirements" for picture targets (Sperber et al., 1979, p. 344). These lower evidence requirements should facilitate early encoding processes, presumably primarily the pictorial-semantic conversion of the pictorial information. Whether picture targets really have lower recognition thresholds in categorically primed conditions, relative to unprimed conditions, is however not directly tested in the study of Sperber at al. (1979), a point that the authors readily acknowledge. The observation that facilitation effects emerge mainly in long-before primed conditions, however, fits well in with their account, because only in these conditions the prime seems to come early enough to affect target processing already at the recognition stage.

Experimental data that support perceptual accounts of facilitatory blocking effects are almost completely lacking at the moment (with one exception, see next paragraph). Motivation for the perceptual notion comes primarily from the observation that semantic facilitation emerges, if at all, only in the initial phase of a block (first cycle). That is, when the set of target items changes and subjects are faced with a new set of target items. Abdel Rahman and Melinger (2007) have pointed out that object identification may be comparatively difficult in this initial cycle of a block, as indicated by the usually slow naming responses in the first cycle as compared to the significantly faster response times in repetition cycles of the same block. In order to reduce the effort to overcome the initial difficulty of object identification, participants may exploit the context. They may use the category cues available in homogeneous blocks to help identify the pictured objects, which partly compensates for the initial recognition difficulty and speeds up recognition times in homogeneous blocks relative to heterogeneous blocks (see also Aristei et al., 2011). The implicit assumption that recognition difficulties will lose impact on naming latencies in repetition cycles of a block seems intuitively plausible. Once participants are acquainted with the set of targets employed in a block, perceptual demands and the time required to recognize the very same objects again in the next cycle markedly decrease, implying that the net impact of perceptual effects on absolute naming latencies should significantly decrease as well.

Data from an EEG study have recently provided evidence for an early, thus probably perceptual locus of semantic facilitation in blocking paradigms. Lin et al. (2022) found that shorter naming latencies in the first cycle of homogeneous blocks correlated with and can even be predicted by enhanced amplitudes in the N1 component, a component that falls in an early, pre-lexical time range often associated with object recognition. The authors further observed that the N1 modulation persisted but was substantially reduced in repetition cycles of homogeneous blocks in which interference dominated. This indicates that context effects on recognition processes will apparently not completely disappear over cycles, but, as speculated by Abdel Rahman and Melinger, they seem to influence naming latencies primarily in the first cycle of a block. Lin et al. (2022) interpreted their findings as evidence in support of perceptual accounts: The larger N1 in the first cycle of homogeneous blocks likely indicates a reduced object recognition effort in homogeneous blocks, relative to heterogeneous blocks. Whether homogeneous blocks really reduce the amount of computation required to achieve object recognition—for example, due to a lowering of the evidence requirements for the targets, as proposed for primed contexts (cf. Sperber et al., 1979)—is, however, not directly tested in Lin et al.'s study. The same holds for their assumption that participants would benefit from homogeneous blocks in a top-down fashion.

In sum, there are some indications in prior semantic priming and blocking studies suggesting that shorter latencies in long-before primed conditions and first within-block cycles may at least in part be due to perceptual recognition effects. However, perceptual explanations of semantic facilitation have, just like post-perceptual ones, the shortcoming that most of their arguments are based on overall picture-naming times which can in principle be influenced by a number of different processes, perceptual and post-perceptual ones. Accepting that overall naming latencies alone help little to definitely clarify whether context manipulations by priming and blocking will affect the ease of object recognition, we chose a different approach in the present study. Instead of examining the time for correct picture naming, we examined the amount of visual information that has to be minimally processed for correct picture naming. If time savings in related contexts would result from lower evidence requirements for the target in the recognition system, as proposed by Sperber et al. (1979), the threshold amount of visual information for object recognition should be lower in related contexts than in unrelated contexts. If, however, no difference in the threshold amounts of visual information will be found, the notion of more efficient entry-level processes for pictures in related contexts compared to unrelated contexts would be seriously challenged. In other words, if we fail to observe a context difference in the perceptual recognition threshold of an object, the claim that semantic facilitation arises post-perceptually in the word production system would receive indirect support.

The Present Study

The goal of the present study is to examine the perceptual demands of object recognition in categorically related and unrelated contexts. The experimental data should help to clarify whether object recognition can be reached earlier, due to less recognition effort, in categorically related contexts than in unrelated contexts.

We conducted a semantic blocking experiment (Experiment 1) and a semantic priming experiment with word primes (Experiment 2). The picture-naming task ("name the object as fast as possible") and context manipulations were kept as in previous blocking and priming experiments, but the method of picture presentation and the dependent variable was changed (see section Methodological Changes: Stimulus Presentation and Dependent Variable) to tap exclusively into perceptual processes.

In both experiments, we contrasted two types of semantic categories, visually similar categories, for example, upper body clothing or birds, and visually dissimilar categories, for example, buildings or vegetables. This category contrast should help to discover the mechanism underlying object recognition processes (see section The Mechanism Underlying More Efficient Recognition Processes: Two Hypotheses and Their Predictions), but it was also included to explore whether visual characteristics of semantic categories generally matter in blocked/primed picture naming.

All pictures presented in the two experiments were black-and-white line drawings of objects. This type of stimulus materials was chosen for the sake of comparability to context effects reported in previous blocking and priming studies measuring naming speed. Most of the above-reviewed studies used simplified line drawings, usually to minimize undesired confounding visual effects of, for example, color, contrast, ease of object separation, background, etc. In word production research, pictures only serve as a means to cue which object name shall be produced. To make these "cues" for the requested name response as clear as possible, it is typically preferred to use simplified depictions of the objects which reduce their visual appearance to just diagnostic shape features.

Methodological Changes: Stimulus Presentation and Dependent Variable

To be able to detect the level of visual detail that is minimally sufficient for recognition (hereafter referred to as the recognition threshold of an object), we changed the method of picture presentation. Rather than presenting objects in single pictures, we presented them in the form of gradually de-blurring image sequences (see Figure 1). Most of the recent theories on visual perception share the assumption that visual information processing temporally progresses from coarse- to fine-grained or global to local information (for a review, see Hegdé, 2008). In our stimuli, we simulated this progression from coarse- to fine-grained visual information in a controlled manner.

The recognition threshold of an object was defined and measured as the de-blurring level that first allowed for correct object naming. The numeric coding of the 19 different blurring levels of a sequence provided the possible values of the dependent variable. Note, the measurement itself did not include any time measurement. To ensure that the dependent variable is insensitive to time effects, each blurring level was shown for 1 s. These long presentation times of each blurring level should make sure that possible post-perceptual context effects, which we cannot rule out, do not influence our dependent variable. According to the literature, context effects could possibly speed up or slow down naming responses by 10–100 ms (typical range of semantic facilitation/interference effects observed in latencies, see review in section Semantically Primed and Blocked Picture Naming: The Method and Typical Findings). With the long presentation times, possible latency effectsof post-perceptual origin of some 10 to 100 ms can cause the critical naming response to be given slightly faster or later at the critical de-blurring level, but not that the response would be given at an earlier de-blurring level (actually not yet sufficient for correct object identification) or a later de-blurring level (actually more than sufficient).

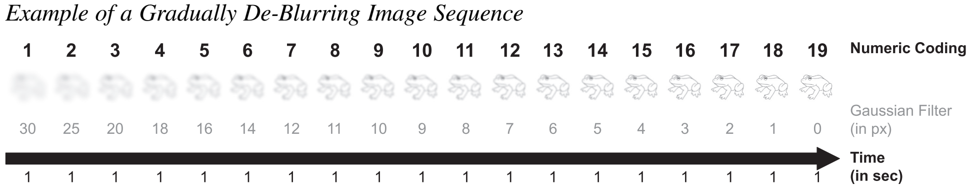
We reasoned that the less visual detail is required for object naming, the less bottom-up analysis would be needed to identify the object and the easier and faster object recognition should be reached. We assumed that processing of visual detail above the critical blurring level is actually unnecessary for tasks demands, suggesting that under the time pressure of conventional single-picture-naming tasks participants might possibly reduce their bottom-up picture processing accordingly to save time and processing effort (see Miller & Bauer, 1981, Experiment 6, for behavioral evidence supporting the assumption that perceptual processing indeed terminates earlier if a more elaborate processing is unnecessary for task demands).

The Mechanism Underlying More Efficient Recognition Processes: Two Hypotheses and Their Predictions

Existing perceptual accounts of semantic facilitation are relatively vague as to the mechanism underlying enhanced object recognition in related contexts. Sperber et al. (1979) proposed for one thing visual priming of object features, presumably a bottom-up mechanism that facilitates purely visual information processing, and secondly, semantically mediated priming, a mechanism that is supposed to facilitate the pictorial-semantic conversion for pictures.

Figure 1

Example of a Gradually De-Blurring Image Sequence



Note. Due to license restrictions of the original pictures the example picture was created for illustration purposes and was not part of the experimental stimuli. Frog image created by authors.

Whether the second mechanism facilitating the pictures' pictorial-semantic conversion is thought to operate bottom-up or top-down is not clear, but Huttenlocher and Kubicek (1983) paraphrased it as "automatic priming" (p. 486) which they expected to occur in "expectancy conditions" (p. 488), suggesting a top-down conception. Lin et al. (2022) also argued for a top-down mechanism, but acknowledged that they cannot specify based on their data whether the mechanism facilitates purely visual information processing or the conceptual interpretation of the visual information, or both.

Taking evidence from object recognition research also into consideration, both the *bottom-up perceptual priming hypothesis* and the *top-down hypothesis* seem reasonable. Neurophysiological studies on visual perceptual priming phenomena have demonstrated that neural responses in the ventral occipitotemporal cortex, an area that is crucial for recognizing visual patterns, gradually decrease when visual information recurs again and again across trials and, accordingly, becomes more familiar (see Wiggs & Martin, 1998, for a review on how the behavioral phenomenon of perceptual priming correlates with repetition suppression). On the other hand, current models of object recognition consider top-down effects an integral part of visual recognition processes, and contexts are assumed to be an important trigger of top-down facilitations (e.g., Bar, 2003; Fenske et al., 2006; Trapp & Bar, 2015). It has been argued that the continuous generation of predictions about the relevant future is the default mode of the brain in order to facilitate the processing of incoming sensory information (e.g., Bar, 2007, 2009; Clark, 2013). The earlier top-down signals are available, the earlier they can "shape and prune ongoing visual processes" (Trapp & Bar, 2015, p. 191), which "significantly reduces the amount of time and computation required for object recognition" (Bar, 2003, p. 601). Top-down facilitations are assumed to result from an enhancement of relevant and a suppression of irrelevant interpretations of the sensory input: not all perceptual object representations stored in memory need to be considered when one tries to find a match that identifies the input, but just a minimal set of the most likely candidates (Bar, 2003, 2004; Bar et al., 2006; Fenske et al., 2006; Trapp & Bar, 2015). The expectations associated with a reduced set of likely candidate interpretations will also guide visual attention (Hochstein & Ahissar, 2002), which "significantly constrains the analysis that needs to be performed by bottom-up processes" (Trapp & Bar, 2015, p. 192), and which usually renders the analysis of really fine-grained information unnecessary for object recognition (Bar, 2003).

To find out whether bottom-up visual priming or top-down influences based on expectancies increase the efficiency of recognition processes in related contexts, we investigated object recognition in pictorial (Experiment 1) and non-pictorial contexts (Experiment 2), and contrasted visually similar and dissimilar categories which differ by definition in the number of visual object features that are shared between-category members.

If more efficient recognition processes would result from bottom-up perceptual priming, a context difference in recognition thresholds should only be found in pictorial contexts (cf. "visual priming should only occur when both the prime and target are pictures," Sperber et al., 1979, p. 344). Thus, lower recognition thresholds in related than in unrelated contexts might be expected to occur in the blocking paradigm (Experiment 1), but should not occur in our priming experiment using word primes (Experiment 2). The bottom-up perceptual priming hypothesis would further predict that the context effect should be larger for visually similar categories than visually dissimilar categories because within-category visual similarity should increase the number of visual features that can be visually primed (cf. the similarity-specific perceptual priming effects found by Flores d'Arcais & Schreuder, 1987).

The alternative top-down hypothesis predicts facilitated object recognition in all types of predictive contexts. Thus, a context difference in recognition thresholds should be obtained in both our experiments. To test whether context effects increase with better target predictability, our priming experiment was designed such that the word primes enable reliable category predictability. In Experiment 2, we used superordinate primes that concretely name the semantic category of the target, whereas, in Experiment 1, predications have to be generated from single exemplars from the same category that appears earlier in homogeneous blocks. As top-down facilitations appear to increase the more accurate and certain expectations are (cf. Bar, 2003), it was hypothesized that context differences in recognition thresholds should be larger in the category priming experiment than in the blocking experiment.

The contrast of visually similar and dissimilar categories should help uncover the type of predictions that may give rise to top-down facilitations. If top-down facilitations are due solely to semantic-taxonomic predictions (i.e., expectations about the semantic concept seen next), which might facilitate the pictorial-semantic conversion for pictures, then the visual type of semantic categories should not affect the magnitude of top-down facilitations in related contexts, as long as the degree of within-category semantic relatedness is equal for visually similar and dissimilar categories. If, however, top-down facilitations are based on perceptual predictions (i.e., expectations about the shape seen next), which might already facilitate the visual processing of pictures, then the visual type of semantic contexts should matter for the magnitude of top-down facilitations. The benefit of perceptual predictions should be large when the expected shape candidates are visually dissimilar, but should be much reduced when the expected and to be discriminated shape candidates look similar (e.g., when some bird has to be identified). In other words, if top-down facilitations would arise from perceptual predictions, then context effects on recognition thresholds should be smaller for visually similar categories than visually dissimilar categories (see Esterman & Yantis, 2010, for fMRI-data and Hirschfeld et al., 2008, for ERP-data that support the notion that top-down recognition effects may arise from perceptual, rather than semantic expectations).

Experiment 1: Semantic Blocking

This experiment investigated whether context manipulation by semantic blocking affects the perceptual recognition threshold of objects. Since perceptual accounts of blocking effects have only been proposed for first within-block cycles, we investigated object recognition thresholds in unrepeated picture presentations only. For this purpose, the non-cyclic version of the blocking paradigm was used.

Method

Participants

Forty-two native German speakers took part in this experiment and were paid for participation. All had normal or corrected-to-normal vision. The data of two participants had to be discarded due to a technical error and a wrong list assignment, respectively. The data of 40 participants (28 female, $M_{\text{age}} = 26.03$) were analyzed.

We determined our sample size based on previous blocking studies reporting facilitatory context effects (e.g., Belke et al., 2017, using $n = 41$). We conducted a sensitivity analysis using G*Power Version 3.1.9.6 (Faul et al., 2009) to estimate the smallest effect size that can be detected with that sample size. The analysis was conducted for our primary question of interest, the context effect (heterogeneous vs. homogeneous blocks), using the parameters $alpha = 0.05$ and $power = 0.8$. The analysis showed that a within-subject study using a sample size of 40 participants should be properly powered to detect context effects of size $d_z = 0.4$ or higher, which is an effect size we considered as sufficiently high to be of practical significance to potentially account for latency effects in first within-block cycles that have been found in prior blocking studies.

Materials

The experimental materials consisted of 60 black-and-white line drawings of common objects, including five objects each from 12 semantic categories (see Figure 1 for a picture example and Appendix for a complete list of items). Six categories were visually similar (VS) categories (tools, upper body clothing, electrical appliances, birds, plants, ungulates) and the other six were visually dissimilar (VD) categories (buildings, tableware, road vehicles, vegetables, insects, fruits). The mean word length of the object names as well as their mean lemma frequency were matched between the two category types as closely as possible (mean number of syllables: 2.2. in VS categories, 2.4. in VD categories; mean normalized (per million) log-transformed frequency: 0.78 for VS categories, 0.74 for VD categories based on the DWDS reference and newspaper corpus, https://www.dwds.de/r?corpus=public). The selection of categories and objects was based on the results of four pretests (see the subsections on Pretest 1–4 below).

All line drawings were normalized with respect to orientation, brightness, color, and image size; complex line drawings were additionally simplified by using Photoshop (Adobe Photoshop CS5 extended, Version 12.0) or Picto-Selector (http://www.pictoselector.eu/). The normalized versions of the line drawings had black lines on white background, horizontal and vertical resolutions of 96 dpi, 32-bit depth, and an image size of 283 px by 283 px.

Gaussian filters of 18 different filter radii (measured as number of pixels) were applied to each normalized line drawing. The resulting 19 versions of each line drawing were composed to an image sequence such that the first picture was an extremely blurred version and the last one the original unfiltered version, see Figure 1. Each image sequence lasted 19 s with an ISI of 0 ms between two successive blurring levels.

Two item groups of 30 image sequences each were formed. Each item group included three VS and three VD categories. The items of each group were combined to form six homogeneous and six heterogeneous blocks. Homogeneous blocks included five objects from one category (e.g., parrot, owl, pigeon, stork, eagle); heterogeneous blocks included one object from each of five different categories of the respective item group (e.g., parrot, pliers, ant, coat, cherry). We controlled that each block included onset-unrelated object names, in one heterogeneous and one homogeneous block it could not be avoided that two object names started with the same onset /t/.

Two lists were compiled. The first list included the six homogeneous blocks of the first item group and the six heterogeneous blocks of the second item group, and vice versa for the second list. That way, each participant saw all items exactly once throughout the experiment either in a homogeneous or a heterogeneous block. A Latin square design was used to assign the lists to participants.

We used a blocked presentation of lists by contexts (i.e., AAAAAABBBBBB with A and B representing the two context conditions). The order of the context conditions was counterbalanced for each list according to a Latin square design. The order of the six blocks of each context part and the order of items within the blocks were randomized for each participant individually.

Twenty-two additional image sequences of other objects were selected as filler items and presented in warm-up trials. Filler objects were one additional object each from the 12 test categories (presented as block-initial item in homogeneous blocks), one object each from six unrelated categories (presented as block-initial item in heterogeneous blocks), two objects from another unrelated category (constituting a homogeneous warm-up block), and two other unrelated objects (constituting a heterogeneous warm-up block).

The experimental items were selected stepwise based on the results of four pretests in which the naming agreement, the visual category types, and the visual and semantic similarity between items were assessed. The selection procedure started out with 330 monochrome line drawings of objects from 25 different semantic categories, which were taken from the IPNP picture database (International Picture Naming Project, https://crl.ucsd.edu/experiments/ipnp/), open-source materials for speech therapists (https://madoo.net/) and teachers (https://www.grundschulmaterial.de/), and a free google search.

All pretests were conducted as web-based studies (using SoSci Survey, https://www.soscisurvey.de/) with undergraduates as participants. All pictures presented in the pretests were normalized and unfiltered pictures. Participants were presented only with a small proportion of the materials (category-balanced lists), so that each pretest lasted about 5 min. All ratings were aggregated using the median.[3]

Pretest 1

Naming agreement. In Pretest 1, participants ($n = 95$) were asked to give a name for the object shown in a picture that first came to their minds. Only pictures with naming agreements above 74% were considered in the remaining pretests. All objects with lower naming agreements and all categories with less than five remaining exemplars were discarded. In total, 131 pictures from 18 different categories were kept and further pretested.

Pretest 2

Classification of the category types. In Pretest 2, pairwise shape similarity ratings for all within-category combinations of the pictures

[3] Although Likert scales provide ordinally scaled data, usually means are calculated for rating experiments. Because of our relatively small number of data points per item (pair) (between 10 and 16), we used the more appropriate median as aggregator in all pretests to prevent outliers to become influential data points.

were collected (in total 446 picture pairs). Picture pairs were presented one after another in random order. Participants ($n = 60$) had to judge the visual similarity for each pair on a Likert scale ranging from 1 (*totally dissimilar shapes*) to 5 (*very similar shapes*). Categories with a median within-category shape similarity of 3 or lower were classified as VD categories, those with a median within-category shape similarity higher than 3 were classified as VS categories. We selected six VS and six VD categories as test categories for the experiment. The number of pictures in each category was reduced to five.

The last two pretests were conducted to assess the semantic and visual similarity within and between categories for the selected 60 test items.

Pretest 3

Semantic similarity. In Pretest 3, semantic similarity ratings for all pairwise picture combinations were collected (in total 1,770 picture pairs). Participants ($n = 46$) were asked to judge the semantic similarity of the two concepts presented in a pair on a Likert scale ranging from 1 (*semantically different/distinct categories*) to 5 (*semantically similar/same category*).[4] The results showed a median semantic similarity of 1 (IQR = 0) for between-category pairs and 5 (IQR = 0) for within-category pairs, verifying a semantic relatedness within the intended categories only. The median within-category semantic similarity was 5 for both VS and VD categories (both IQRs = 0), indicating that the degree of semantic relatedness was equal for both types of categories (for category-specific semantic similarities values, see Appendix).

Pretest 4

Visual similarity[5]: The procedure of Pretest 4 was the same as in Pretest 3, except that participants ($n = 62$) were asked to judge the visual similarity between the objects of a pair. A Likert scale ranging from 1 (*totally dissimilar shapes*) to 5 (*very similar shapes*) was used. The results showed that the median visual similarity of objects from different categories was low (*Mdn* = 1, IQR = 1). The median within-category visual similarity was 2 for VD categories (IQR = 2) and 5 for VS categories (IQR = 1; for category-specific visual similarities values, see Appendix). The difference in the within-category visual similarity between VS and VD categories was statistically significant, $\chi^2(1) = 29.9$, $p < .001$.[6]

Design

A two-factorial design was used, including the independent variables: Context (Blocking) (categorically homogeneous vs. categorically heterogeneous) and Category Type (VS vs. VD). Context (Blocking) was manipulated within subjects and within items; Category Type was a within-subjects, but between-items variable.

As described in the Introduction, the dependent variable was the de-blurring level that first allowed for correct object naming, measured in terms of the position of the blurring levels in the image sequence (see Figure 1). Thus, the measured variable can take on values between 1 (extremely blurred level) and 19 (unblurred level).

Procedure

Participants were tested individually in a dimly lit sound-insulated room. The distance to the computer screen was approximately 60 cm.

Prior to the experiment, participants received a written instruction about the procedure and task. There was no familiarization neither with the pictures nor the expected names. The participants were asked to name the object they believed to see as fast and as specifically as possible directly during the image sequences by using a bare noun (e.g., *Papagei* "parrot"). Specifications or corrections during a sequence were encouraged.

The experiment proper was divided into two parts, one for each context condition. Between the two parts, participants were offered a break in which they could leave the room.

Each part started with a warm-up block of two image sequences. The warm-up block was categorically homogeneous, if homogeneous blocks were presented in the upcoming part, and categorically heterogeneous, if heterogeneous blocks were presented in the upcoming part. Then, the six experimental blocks of each part were carried out. The first sequence of an experimental block was always a filler sequence presenting a further category member of the target category in homogenous blocks or an unrelated item in heterogeneous blocks. Participants started each block individually per button press. Sequences within a block started automatically.

The trial structure was as follows. A fixation cross appeared at the center of the screen for 500 ms followed by a blank period of 200 ms. Then, the image sequence was shown centrally. A trial finished with a blank intertrial interval of 1,000 ms. The onset of an image sequence was auditorily signaled by a sound of a short gong; the subsequent picture onsets within the sequence were accompanied by a tick-sound. The sounds were included to attract the participants' attention constantly and have a signal coding of the picture onsets in the recordings.

Overall, 76 image sequences were presented in each testing session (60 experimental sequences and 16 filler sequences). The complete session lasted about 30 min.

Apparatus

The experiment was controlled by Presentation Software (Version 18.3). The stimuli were presented on a BenQ-monitor (XL2430) with a screen resolution of 1920 px by 1080 px and a vertical refresh rate of 60 Hz. We used a black screen background for stimulus presentation.

Naming responses were registered by using a Sennheiser-headset with attached microphone. All responses were simultaneously

[4] The original coding was 1 (*semantically similar/same category*) to 5 (*semantically different/distinct categories*). The scale and results are reported with the inverted numerical coding for better comparability with the visual similarity data.

[5] The data from Pretest 4 were also used for another study that is published in Bechberger and Scheibel (2020).

[6] The difference between the medians was statistically analyzed with Cumulative Link Mixed Model in R using the *clmm* function from the package *ordinal* (Version 2019.12-10; https://CRAN.R-project.org/package=ordinal). *p* values were estimated by a likelihood ratio test comparing the model including the variable Category Type as fixed effect against an intercept only model. The random effects structure included random intercepts for the specific categories.

recorded as WAV files on a Marantz professional recorder (PMD 620). The recorded WAV files were manually analyzed in Praat (Version 6.1.08; http://www.praat.org/).

Data Preparation

We categorized all responses given in the course of an image sequence. Expected target names (e.g., parrot in a parrot trial) were categorized as target naming. All other responses counted as misnaming and were categorized with respect to their type of misnaming, for instance as a synonym (e.g., budgerigar in a parrot trial), coordinate (e.g., owl in a parrot trial), or category name (e.g., bird in a parrot trial). The first response with the expected target name and no false revisions within the remaining time of a sequence was defined as the critical target response. Later responses with target repetitions or alternative target names from the same or lower taxonomic levels (synonyms or hyponyms) did not count as false revisions. The picture number visible during the critical target response counted as the critical picture number of the trial. The critical picture number was adjusted backwards, if the time-window between the onset of the picture and the target response was shorter than 500 ms.

For an explorative analysis, we also determined the critical picture number of first-category recognition. We defined the first response related to the target category (category naming, target naming or hypernyms, hyponyms, synonyms or names of coordinates) and no false category revisions thereafter as the critical naming indicating category recognition. Response repetitions or responses with other object names also related to the target category did not count as false category revisions. The critical picture number for category recognition was again adjusted backwards for latencies shorter than 500 ms, otherwise, the picture number visible during the critical category response counted as the critical picture number.

Statistical Analysis

The data were analyzed with R and the packages *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017). We fitted a linear mixed-effects model (LMM) that included the two variables from our experimental design Context (Blocking) and Category Type as fixed effects with interaction terms. Both variables were contrast coded with −1 and 1. The random effects structure included random intercepts for participants and items as well as by-participant and by-item random slopes for the within-subject and within-item variable Context (Blocking).[7] The built-in optimizer *bobyqa* and a higher number of iterations were used to obtain convergence of the model. A visual inspection of residual plots did not reveal critical deviations from normality, linearity, and homoscedasticity.

Transparency and Openness

The design of Experiment 1 and its analysis were not preregistered. A complete list of stimuli, ratings collected in Pretest 3 (semantic similarity) and Pretest 4 (visual similarity) as well as the data collected in Experiment 1 and the analysis code being used are publicly available in OSF at https://osf.io/r5kyp/?view_only=d6c19885fcee447483205c53071af5d7. The stimulus pictures cannot be made publicly available due to license restrictions of the original pictures but will be provided upon request.

Results

Trials without a target name response amounted to 9.5%; trials without a target category response amounted to 0.63%. The data with missing target responses were excluded from the respective analyses.

Target Naming

The descriptive results are shown in Figure 2 as a function of Context (Blocking) and Category Type. A significant main effect of Context (Blocking) reflected that the critical picture number was lower in homogeneous blocks ($M = 9.35$, $SD = 4.06$) compared to heterogeneous blocks ($M = 10.02$, $SD = 3.97$; effect size $d_z = 0.75$). A significant main effect of Category Type reflected that the critical picture number of objects of VD categories ($M = 8.78$, $SD = 4$) was lower compared to that of the object of VS categories ($M = 10.63$, $SD = 3.83$; effect size $d_z = 2.05$). Furthermore, the interaction of Context and Category Type was statistically significant (see Table 1), reflecting that the effect of homogeneous contexts was more pronounced for objects of VD categories than it was for objects of VS categories (see Figure 2B).

Category Recognition

Figure 3 displays from which blurring level onwards the target category was correctly identified. For VS categories, the critical picture number of first category recognition was similar in homogeneous ($M = 8.85$, $SD = 3.65$) and heterogeneous ($M = 8.86$, $SD = 3.49$) blocks. For VD categories, by contrast, the critical picture number in homogeneous blocks was about one and a half levels lower ($M = 7.49$, $SD = 3.49$) compared to heterogeneous blocks ($M = 8.78$, $SD = 3.87$, see Figure 3B). A significant interaction between Context and Category Type confirmed this differential effect of Context on VD compared to VS categories (see Table 2).

Discussion

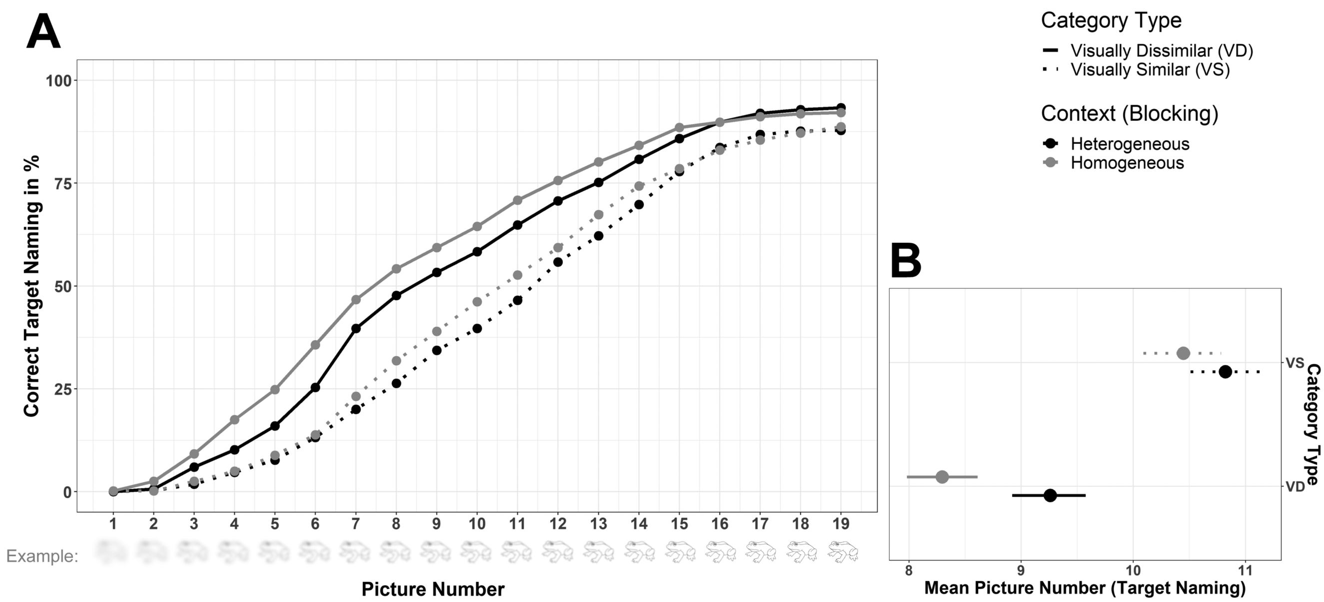
The findings of Experiment 1 can be summarized as follows: homogeneous blocks lowered the amount of visual information minimally required for object recognition, relative to heterogeneous blocks, with visually dissimilar (VD) categories benefiting more from homogeneous blocks than visually similar (VS) categories did.

The significant context effect confirms that less visual analysis and encoding suffice to identify depicted objects in homogeneous blocks as compared to heterogeneous blocks. This provides direct evidence for the prediction of perceptual accounts of facilitatory blocking effects that object recognition should be easier in homogeneous blocks than in heterogeneous blocks (Abdel Rahman & Melinger, 2007; Aristei et al., 2011; Lin et al., 2022). The relatively lower recognition threshold of objects in homogeneous blocks indicates that there is indeed a possibility to save time during recognition processes in homogeneous blocks that does not exist in a comparable way in heterogeneous blocks.

The observed category type effect indicates that the visual type of semantic categories generally matters in blocked picture naming, because it influences by itself the difficulty of object recognition.

[7] A random slope for the interaction effect of Context (Blocking) and Category Type was initially included for participants, but removed to avoid singular fit of the model.

Figure 2
Results of Experiment 1 (Target Naming)



Note. (Panel A) Mean percentage of correct target naming at each picture number of the image sequence. (Panel B) Mean critical picture numbers of first target naming with 95% bootstrapped CIs. Due to license restrictions of the original pictures the example picture was created for illustration purposes and was not part of the experimental stimuli. Frog image created by authors.

A semantic confound of this category type effect can be ruled out because we controlled that within-category semantic relatedness was equally high for VS and VD categories (see Pretest 3, Appendix for category-specific values). The direction of the category type effects, indicating higher perceptual demands for objects of VS than VD categories, is in line with earlier findings of visual similarity effects on visual perception (e.g., Ashworth III & Dror, 2000; Humphreys et al., 1988; Laws & Gale, 2002). The fact that our paradigm was able to replicate an established effect on visual perception confirms that our novel dependent variable is sensitive to perceptual recognition processes.

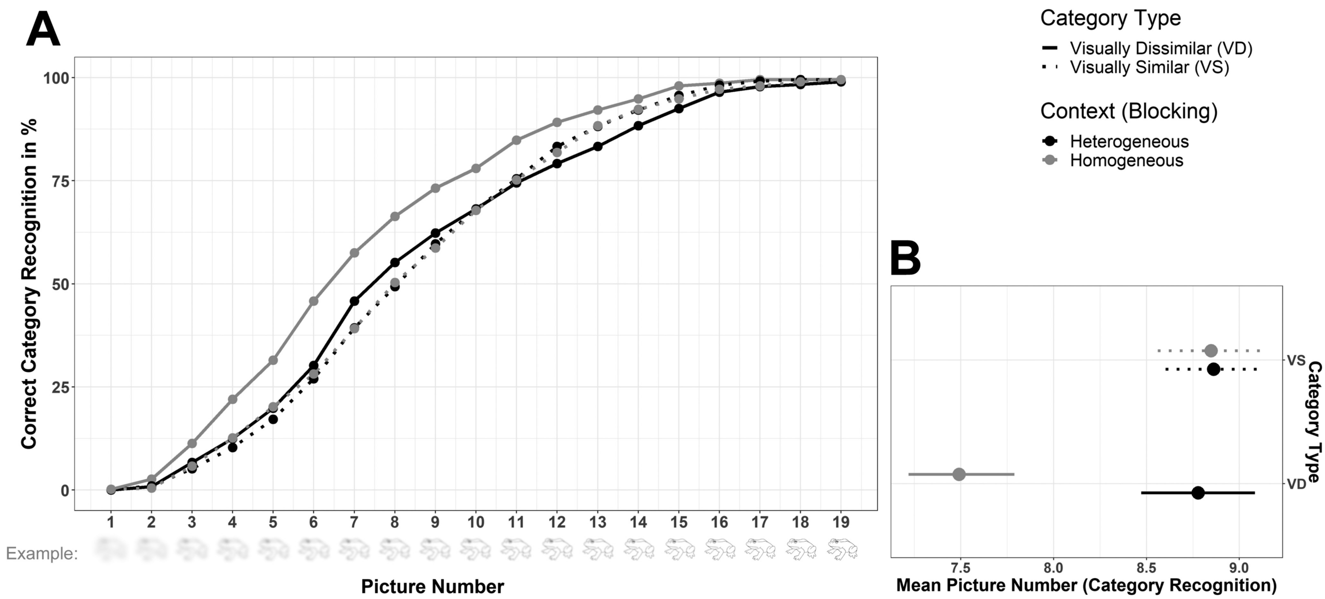
The obtained interaction of context and visual type of semantic categories demonstrated that context effects are influenced by visual characteristics of categories that go beyond their mere semantic-taxonomic definition. This finding challenges all hypotheses that suggest a purely semantically driven mechanism underlying context effects, and is, in

Table 1
Analysis (LMM) of Experiment 1 (Target Naming)

Fixed effects					
Predictors	Estimates	*SE*	*df*	*t*	*p*
(Intercept)	9.80922	.36533	81.72658	26.850	<.001***
ContextBlocking Het—Hom	0.33313	.06931	28.47538	4.806	<.001***
CategoryType VD—VS	0.93232	.32971	59.99393	2.828	<.01**
CategoryType:ContextBlocking	−0.14070	.06751	55.24502	−2.084	.042*
Random effects					
Groups	Variance	*SD*			
SubjID					
(Intercept)	0.99029	0.9951			
Het—Hom	0.01000	0.1000			
ItemID					
(Intercept)	6.29990	2.5100			
Het—Hom	0.05345	0.2312			
Residual	7.89540	2.8099			
N Observations: 2,172; *N* Subjects: 40; *N* Items: 60					

Note. *SE* = standard error; *SD* = standard deviation; LMM = linear mixed-effects model. Formula of the model: CriticalPictureNumber_Target ~ CategoryType × ContextBlocking + (ContextBlocking | SubjID) + (ContextBlocking | ItemID), including the contrasts Het—Hom (heterogeneous vs. homogeneous blocks) and VD—VS (visually dissimilar vs. visually similar categories).
*$p < .05$. **$p < .01$. ***$p < .001$.

Figure 3
First Target Category Recognition (Experiment 1)



Note. (Panel A) Mean percentage of correct category recognition at each picture number of the image sequence. (Panel B) Mean critical picture numbers of first category recognition with 95% bootstrapped CIs. Due to license restrictions of the original pictures the example picture was created for illustration purposes and was not part of the experimental stimuli. Frog image created by authors.

particular, inconsistent with the predictions of the semantic top-down hypothesis assuming semantic-taxonomic expectations to give rise to facilitated object recognition in related contexts.

An interaction of context and category type was predicted by the perceptual priming hypothesis, but the direction of the interaction that we have found is inconsistent with what was expected by the bottom-up priming hypothesis. The results showed that sharing a high number of shape features with previous items in related contexts was not advantageous, as predicted by the perceptual priming hypothesis, but attenuated facilitatory blocking effects. This casts doubt on bottom-up visual priming as the underlying mechanism of contextual facilitations. A second finding that is difficult to reconcile with bottom-up priming is the observation that there was a context effect for VD categories given the test materials we used. The results of Pretest 4 attested that the degree of visual similarity within-categories and between-categories was low for VD categories, and crucially: almost equal (see

Table 2
Analysis (LMM) of First Target Category Recognition (Experiment 1)

Fixed effects					
Predictors	Estimates	*SE*	*df*	*t*	*p*
(Intercept)	8.49736	.33453	93.47351	25.401	<.001***
ContextBlocking Het—Hom	0.32521	.10211	57.72505	3.185	<.01**
CategoryType VD—VS	0.35805	.27258	59.90697	1.314	>.1
CategoryType:ContextBlocking	−0.32178	.08327	59.26781	−3.864	<.001***
Random effects					
Groups	Variance	*SD*			
SubjID					
(Intercept)	1.5046	1.2266			
Het—Hom	0.1397	0.3737			
ItemID					
(Intercept)	4.2815	2.0692			
Het—Hom	0.2397	0.4895			
Residual	7.0097	2.6476			
N Observations: 2,385; *N* Subjects: 40; *N* Items: 60					

Note. *SE* = standard error; *SD* = standard deviation; LMM = linear mixed-effects model. Formular of the model: CriticalPictureNumber_Category ∼ CategoryType × ContextBlocking + (ContextBlocking | SubjID) + (ContextBlocking | ItemID), including the contrasts Het—Hom (heterogeneous vs. homogeneous blocks) and VD—VS (visually dissimilar vs. visually similar categories).
$^{**}p < .01$. $^{***}p < .001$.

Appendix). Thus, for VD categories, the items in both homogeneous and heterogeneous blocks had only few visual features in common, raising the question of how bottom-up priming should have been possible for VD categories, and why it could occur in one visually dissimilar context but not the other.

The results obtained in Experiment 1 are entirely consistent with the predictions of the perceptual top-down hypothesis. This corroborates the assumption that perceptual predictions, that is, top-down expectations about the shape seen next, play a major role in facilitatory blocking effects on recognition processes.

The explorative analysis yielded interesting observations, the interpretations of which however must remain tentative because we analyzed first category(-related) naming although there was no category naming task. A comparison of the thresholds of category recognition (Figure 3B) and target recognition (Figure 2B) indicated that objects were first identified at a category level: the blurring level that allowed for category identification was generally too impoverished for naming the concrete category exemplar, in both related and unrelated contexts and for both VS and VD categories. The category-before-object finding is in line with similar findings reported in recent recognition studies (see e.g., Clarke et al., 2013, 2015; Gauthier et al., 1997; Grill-Spector & Kanwisher, 2005; Poncet & Fabre-Thorpe, 2014; Thorpe et al., 1996; VanRullen & Thorpe, 2001). Taking a closer look at the category recognition data alone (Figure 3B), a context effect similar to that observed in the target naming data can be observed, at least for VD categories. This suggests that related contexts may facilitate both initial broad categorizations of the visual input and the identification of the concrete object. Importantly, even in homogeneous blocks category names were not produced for the first or second blurring level shown in an image sequence. This indicates that participants did not simply guess the category; they rather seem to have waited until a certain level of sensory input confirmed a particular category.

In sum, the results of Experiment 1 are most consistent with the predictions of the perceptual top-down hypothesis. However, due to the pictorial format of blocking contexts, it cannot be definitely ruled out that bottom-up priming of shape features may have contributed to the obtained context effects. To test whether facilitations on object recognition in related contexts hinge on a pictorial presentation of the contexts, we conducted a second experiment. In Experiment 2, we changed the context modality from pictorial to non-pictorial.

Experiment 2: Category Priming

This experiment tested whether context manipulation by category priming with superordinate word primes affects the perceptual recognition threshold of objects. The same set of test materials as in Experiment 1 was used. As in Experiment 1, participants saw only one image sequence per object.

Method

Participants

Thirty-seven (including one participant who needed to be replaced due to technical errors in the recording) native German speakers not participating in Experiment 1 took part in this experiment. Participants were paid for participation. All had normal or corrected-to-normal vision. The data of 36 participants (28 female, $M_{age} = 22.81$) were analyzed.

The minimum number of participants required by our design and the planned counterbalancing was 36. We conducted a sensitivity analysis using G*Power, Version 3.1.9.6. (Faul et al., 2009) to estimate the smallest effect size that can be detected with that sample size. The analysis was conducted for our primary question of interest, the context effect (heterogeneous vs. homogeneous blocks), using the parameters $alpha = 0.05$ and $power = 0.8$. The analysis showed that a within-subject study testing 36 participants should be properly powered to detect context effects of size $d_z = 0.42$ or higher. Given the effect size of the context effect observed in Experiment 1 ($d_z = 0.75$), we assumed that our sample size of 36 participants should be large enough to detect context effects, even when it turns out that the effect of word contexts is smaller than the effect of pictorial contexts, in which perceptual priming is (additionally) possible.

Materials

The test materials were the same as in Experiment 1, 60 gradually de-blurring image sequences of common objects (see Appendix for a complete list of the test items), plus 14 fillers (a subset of the fillers used in Experiment 1).

The image sequences were combined to form six blocks of 10 sequences each. In half of the blocks, a category prime was shown before the target picture, in the other half no prime stimulus preceded the target pictures. To avoid grouping effects, three different grouping versions of the items were created; each version was tested with the same number of participants.

All blocks were categorically heterogeneous (each object belonged to a distinct category). Each block consisted of five objects from VS categories and five objects from VD categories; semantically associated items within blocks were avoided. We controlled that all object names of a block had different phonetic onsets. The order of blocks was counterbalanced across participants using a Latin square design. The order of items within each block was pseudo-randomized for each participant individually, such that maximally three items of the same category type (VS or VD), and maximally three natural or artificial items appeared in successive trials. Each participant saw each item exactly once.

As in Experiment 1, we used a blocked presentation of lists by context (i.e., AAABBB with A and B representing the two context conditions). The order of context conditions was counterbalanced for each block order of each list.

German category names served as superordinate primes. The category names were presented centrally on the screen in white letters, Arial font, 30 pt. To assess the category names most common for the selected objects and well familiar to our test population, we conducted a web-based survey (https://www.soscisurvey.de/) with 46 participants in advance. Category pictures, including the five test items of a category, were presented, interspersed with filler pictures showing other categories. Participants were asked to give a superordinate category name for each category picture. The names most frequently mentioned were used as category names (see Appendix).

Design

As in Experiment 1, a two-factorial design was used with the independent variables Context (Priming) (category prime vs. no prime) and Category Type (VS vs. VD). Context (Priming) was manipulated within subjects and within items; Category Type was a within-subjects and between-items variable. The same dependent variable as in Experiment 1 was used, that is, the de-blurring level that first allowed for correct object naming measured in terms of the position of the blurring levels in the image sequence (see Figure 1).

Procedure

The laboratory setting and the task were the same as in Experiment 1, including no familiarizations of any kind. As in Experiment 1, the experiment proper included two parts, one for each context condition. Between the two parts, participants were offered a break.

Each part started with a warm-up block of four image sequences. The warm-up block was of the same context condition as the upcoming experimental part. The first image sequence of each experimental block was always a filler sequence showing an object from an unrelated category. Participants started each block individually per button press. Image sequences within a block started automatically.

The trial structure was as follows. A fixation cross was presented at the center of the screen for 500 ms followed by a blank period of 200 ms. Then, in the condition with category primes, the name of the target category was presented centrally for 1,000 ms followed by the image sequence of the target. In the condition without primes, the image sequence followed the blank screen immediately. Each trial finished with a blank intertrial interval of 1,000 ms. As in Experiment 1, the onset of an image sequence was auditorily signaled by the sound of a short gong; the subsequent picture onsets within the sequence were accompanied by a tick-sound.

Overall, 74 image sequences were presented in each testing session (60 experimental sequences and 14 filler sequences). The complete session lasted about 30 min.

Apparatus, Data Preparation, and Statistical Analysis

The apparatus and the method of data preparation were the same as in Experiment 1.

As in Experiment 1, we fitted a linear mixed-effects model (LMM) that included the two variables from our experimental design Context (Priming) and Category Type as fixed effects with an interaction term. Both variables were contrast-coded with −1 and 1. The random effects structure included random intercepts for participants and items as well as by-participant and by-item random slopes for the within-subject and within-item variable Context (Priming). All model optimizations used in Experiment 1 were likewise necessary in order to obtain convergence or avoid singular fit of the models.

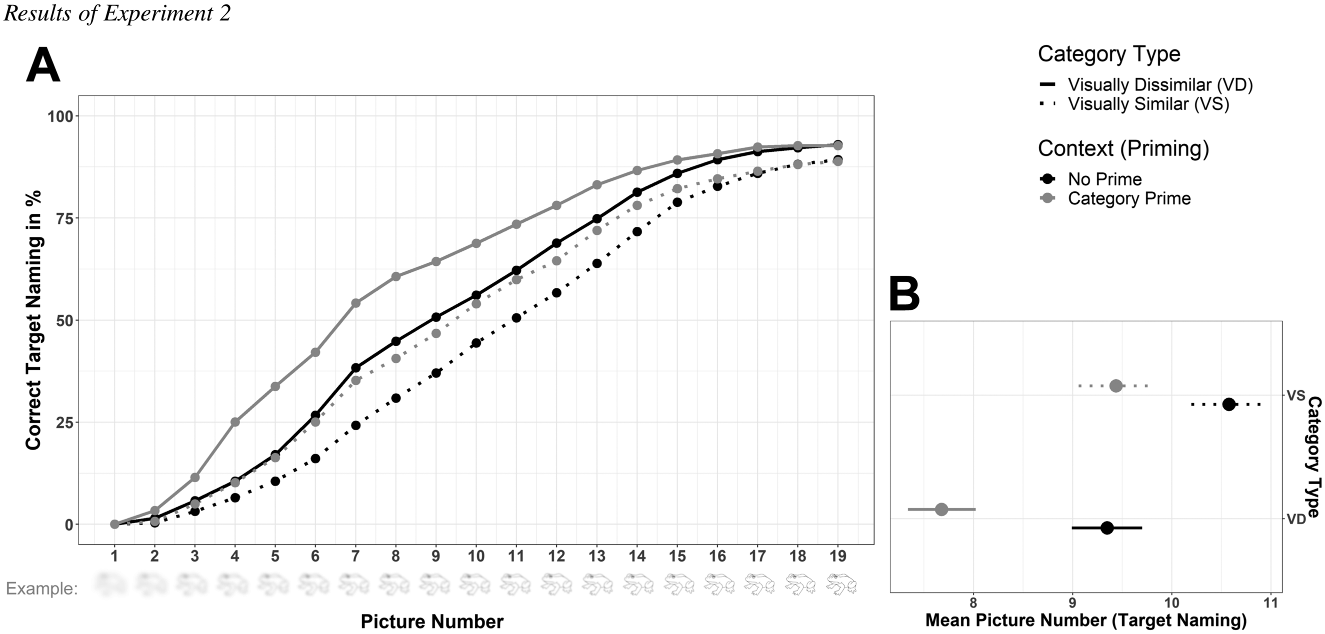
Transparency and Openness

The design of Experiment 2 and its analysis were not preregistered. A complete list of stimuli, the data collected in Experiment 2, and the analysis code being used are publicly available in OSF at https://osf.io/r5kyp/?view_only=d6c19885fcee447483205c53071af5d7. The stimulus pictures cannot be made publicly available due to license restrictions of the original pictures but will be provided upon request.

Results

Experimental trials without a correct target name response (9.12%) were excluded from the analysis. Figure 4 shows the descriptive results as a function of Context (Priming) and Category Type. On average, the critical picture number was nearly one and a half levels lower in primed contexts ($M = 8.54$, $SD = 4.1$) compared to unprimed contexts ($M = 9.95$, $SD = 4.11$; effect size $d_z = 1.12$). Objects of VD categories ($M = 8.51$, $SD = 4.09$) had critical picture

Figure 4
Results of Experiment 2



Note. (Panel A) Mean percentage of correct target naming at each picture number of the image sequence. (Panel B) Mean critical picture numbers of first target naming with 95% bootstrapped CIs. Due to license restrictions of the original pictures the example picture was created for illustration purposes and was not part of the experimental stimuli. Frog image created by authors.

Table 3
Analysis (LMM) of Experiment 2

Fixed effects					
Predictors	Estimates	*SE*	*df*	*t*	*p*
(Intercept)	9.38899	.37842	81.83552	24.811	<.001***
ContextPriming No—Prime	0.70633	.08528	33.71958	8.283	<.001***
CategoryType VD—VS	0.82497	.33841	60.02132	2.438	.018*
CategoryType:ContextPriming	−0.14347	.07737	58.00354	−1.855	.069
Random effects					
Groups	Variance	*SD*			
SubjID					
(Intercept)	1.03195	1.0159			
No—Prime	0.04628	0.2151			
ItemID					
(Intercept)	6.60485	2.5700			
No—Prime	0.09637	0.3104			
Residual	8.51987	2.9189			
N Observations: 1,963; *N* Subjects: 36; *N* Items: 60					

Note. *SE* = standard error; *SD* = standard deviation; LMM = linear mixed-effects model. Formular of the model: CriticalPictureNumber_Target ~ CategoryType × ContextPriming + (ContextPriming | SubjID) + (ContextPriming | ItemID), including the contrasts No—Prime (no prime vs. category prime) and VD—VS (visually dissimilar vs. visually similar categories).
$^{*}p < .05$. $^{***}p < .001$.

numbers which were one and a half levels lower on average compared to those of VS categories ($M = 10.01$, $SD = 4.1$; effect size $d_z = 1.7$). Both main effects were statistically significant (see Table 3). Numerically, primed contexts lowered the critical picture numbers more strongly for VD as compared to VS categories (see Figure 4B), however, the interaction between Context and Category Type was only marginally significant (see Table 3).

Discussion

The results of Experiment 2 extend those of Experiment 1 in that they show a facilitatory effect of related contexts even in non-pictorial contexts. Superordinate word primes before the target lowered the amount of visual information minimally required for object recognition, relative to unprimed conditions. As in Experiment 1, this context difference in recognition thresholds was numerically larger for visually dissimilar (VD) categories than visually similar (VS) categories.

The replication of the context effect in a priming paradigm using word primes indicates that increased recognition efficiency in categorically related contexts does not hinge on bottom-up visual priming. Words do not share visual features with pictorial representations of objects in any systematic way (Sperber et al., 1979, p. 344). Thus, the prime stimulus should not have been able to prime bottom-up visual features of related targets in Experiment 2. The fact that we nevertheless observed enhanced object recognition, relative to unprimed conditions, suggests that there may be another mechanism underlying the increased recognition efficiency in related contexts. The alternative assumption that superordinate primes may have given rise to indirect perceptual priming based on mental imagery, that is, a mental visualization of category-typical shape features triggered by the category labels (cf. Zwaan et al., 2002), can be doubted. The finding that context effects were, again, numerically largest for VD categories conflicts with the predictions of the priming hypothesis. Visual priming effects, irrespective of whether directly or indirectly triggered, should be largest for visually homogeneous categories whose members share many features in common. We, however, observed the opposite in both our experiments: categories with visually dissimilar category members benefited the most from related contexts. Thus, as both the context effect obtained in Experiment 2 and the direction of the observed interaction of context and category type are difficult to reconcile with visual priming, our data cast doubt on the assumption that recognition effects in related contexts would arise due to bottom-up perceptual priming.

By contrast, the results obtained in Experiment 2 are, again, entirely consistent with the predictions of the perceptual top-down hypothesis. This further supports the assumption that increased recognition efficiency in related contexts may be due to perceptual predictions, that is, top-down expectations about the shape seen next.

In the next section, the data of Experiment 1 and 2 will be jointly analyzed. The top-down hypothesis predicted that context effects should be larger in the priming experiment than in the blocking experiment because concrete category labels should provide more certain and accurate cues to the target than single pictures of objects were able to do. To test this prediction of the top-down hypothesis, we compared the relative advantage of superordinate word primes with that of homogeneous blocks. It is interesting to note that the object recognition thresholds in the unrelated contexts, that is, heterogeneous blocks and unprimed contexts, were nearly identical in the two experiments (cf. Figure 2B vs. 4B). Thus, it does not appear that the context modality, pictorial or non-pictorial, had any independent effect on recognition processes.

Joint Analysis of Experiments 1 and 2

This analysis was conducted to test whether prior category words were more effective in reducing the object recognition threshold than categorically homogeneous picture blocks were.

Data Preparation and Statistical Analysis

Both experiments employed exactly the same items. Therefore, we determined the magnitude of blocking and priming effects per item and analyzed these item-specific values. For each data set, the difference between the item-specific picture numbers from the unrelated and the related contexts were calculated. These item-specific measures for the relative advantage of priming and the relative advantage of blocking were jointly analyzed in an LMM. The factors Context Type (blocking vs. priming) and Category Type (VD vs. VS) were specified as fixed effects; items were specified as random effect. The interaction of Context Type and Category Type was not included as fixed effect because it did not reach significance ($p > .5$).

Results and Discussion

The joint analysis revealed significant main effects of both Context Type and Category Type (see Table 4).

The direction of the context type effect indicates that the relative advantage of related contexts was stronger (twice as large) in the word-priming experiment (difference primed vs. unprimed condition: $M = 1.38$, $SD = 1.29$) than in the blocking experiment (difference homogeneous vs. heterogeneous blocks: $M = 0.67$, $SD = 1.13$). The detection of a significant context-type effect is important for the top-down hypothesis: It confirms that context effects on recognition processes indeed increase with better target predictability. Single objects seen in the first trials of homogeneous blocks can be associated with various categories and category levels. This uncertainty about the exact category that items in homogeneous blocks will share might have made top-down expectations vaguer and in consequence less effective in the blocking paradigm as compared to our priming paradigm in which primes provided reliable a priori category information about the target. The significant context type effect thus provides further support for the claim that top-down expectations lower the recognition threshold of objects in related contexts. By contrast, the bottom-up priming hypothesis is again challenged. The context type effect indicates that context effects on recognition processes not only do not hinge on a pictorial presentation of contexts (see Experiment 2), but they can even benefit from a non-pictorial presentation under certain conditions. Thus, Sperber and colleagues' (1979, cf. p. 344) hypothesis that target pictures may have the lowest evidence requirements when following the processing of pictorial contexts cannot be confirmed, suggesting, accordingly, that bottom-up visual priming may contribute little to the relative lowering of object recognition thresholds in related contexts.

Table 4
Results of the Joint Analysis of Experiments 1 and 2 (LMM)

Fixed effects					
Predictors	Estimates	*SE*	*df*	*t*	*p*
(Intercept)	1.02453	.11514	60	8.898	<.001***
ContextType Blocking—Priming	−0.35736	.09576	60	−3.732	<.001***
CategoryType VD—VS	−0.31886	.11514	60	−2.769	<.01**
Random effects					
Groups	Variance	*SD*			
ItemID (Intercept)	0.2452	0.4951			
Residual	1.1005	1.0490			
N Observations: 120; *N* Items: 60					

Note. SE = standard error; SD = standard deviation; LMM = linear mixed-effects model. Formular of the model: ContextDiff ~ CategoryType + ContextType + (1 | ItemID), including the contrasts Blocking—Priming (Blocking (Exp1) vs. Priming (Exp2)) and VD—VS (visually dissimilar vs. visually similar categories).
$^{**}p < .01$. $^{***}p < .001$.

Please note, that the variable Context Type was unavoidably confounded by format and timing differences in how and how long category information about the target was available before the target could be named. This may raise the question whether these differences and post-perceptual effects favored by these differences could have given rise to the obtained context-type effect. Timing differences have been shown to be an important factor influencing short-lived semantic priming, which can speed up response times to the target immediately following the prime by up to some hundred milliseconds. In our paradigm, we presented each blurring level shown in an image sequence for 1 s, precisely to prevent such short-lived semantic priming effects from becoming influential (see section Methodological Changes: Stimulus Presentation and Dependent Variable). It is also important to remember that it was never the picture directly following the context stimulus that allowed for the critical target naming response we measured. In both experiments, the first recognizable target picture was on average the eighth to 11th picture shown in an image sequence, implying that there was a sequence of several pictures between the context stimulus and the critical target picture and, more important, the context-target interval was several seconds long (at least about 9 s) in both experiments. This suggests that relatively longer (blocking paradigm) or shorter (priming paradigm) availability of a priori category information might have been of little consequence, and short-lived semantic priming effects did not seem to be an issue in either experiment. Top-down effects could possibly have been boosted by a relatively earlier availability of category cues. This difference could have provided an explanation for potential larger facilitations in the blocking paradigm. We, however, obtained larger facilitations in the priming paradigm, suggesting that the difference in the timing of when category information became available does not account for our context-type effect.

With respect to the format of the contextual cues, a possible concern raised by a reviewer might be that the context in Experiment 2 but not in Experiment 1 was a word. The linguistic input could possibly have had a priming effect at the lexical level which facilitated lexical access in the priming paradigm, relative to the blocking paradigm. Indeed, it is conceivable that processing of category names gives rise to automatic co-activations in the respective semantic network, but in this case, the lexical representation of all category members, rather than solely the target name should benefit from activation spreading. Thus, a cohort of inter-related lexical candidates, rather than exclusively the target name, should become pre-activated by the category names, implying that lexical competition effects may be equally possible. A study that we know of that investigated the effect of category names on the lexical access of targets (a PWI study presenting category names simultaneously with the target, Kuipers et al., 2006) found semantic inhibition of about 50 ms in naming latencies, demonstrating that inhibitory effects of superordinate words apparently outweigh possible priming effects at the lexical level. This makes it unlikely that lexical effects were responsible for the relatively larger facilitations in the word-priming paradigm, relative to the blocking paradigm. Please also note, that time effects associated with lexical-semantic priming/interference should generally not have any impact on our data as we did not measure naming latencies, and presented each blurring level for 1 s. The long presentation times of each blurring level in our paradigm made sure that time effects of some 50–100 ms could possibly cause the critical naming response to be given slightly faster or later at the critical de-blurring level but not that it would have been given at an earlier or later de-blurring level. In sum, this suggests that the difference in the format of category cues may also not account for our context type effect.

The direction of the second main effect, the category type effect, indicates that VD categories benefited more (nearly twice as much) from related contexts (difference related vs. unrelated contexts: $M = 1.34$, $SD = 1.22$) than VS categories did (difference related vs. unrelated contexts: $M = 0.71$, $SD = 1.23$). As shown by the absence of interaction between our two factors, this processing advantage for VD categories occurred in all context modalities, pictorial and non-pictorial contexts. The finding that a high number of visual features shared between-category members generally reduced, rather than increased, the relative advantage of categorically related contexts is problematic for all bottom-up accounts of semantic context effects. It is also problematic for accounts assuming that purely semantic-taxonomic predictions would give rise to contextual facilitations, as those accounts would expect no effect of the visual type of categories. The degree of within-category semantic relatedness was comparable for VS and VD categories in our study (see Pretest 3), suggesting that the benefit of top-down expectations on a purely semantic-taxonomic level should have also been comparable. The only hypothesis that predicted the category-type effect that we found is the perceptual top-down hypothesis. This hypothesis predicted larger facilitations for VD categories because the expectation of a set of highly distinctive shape candidates substantially simplifies object discrimination, whereas the expectation of a set of very similarly looking candidates still requires a rather fine-grained matching procedure to determine which of these candidates the input matches best.

General Discussion

The aim of the present study was to investigate to what extent context manipulations by semantic blocking and semantic priming alter the perceptual demands of object recognition. Perceptual accounts of behavioral blocking/priming effects propose that shorter naming latencies in the first cycle of homogeneous blocks, relative to heterogeneous blocks (e.g., Abdel Rahman & Melinger, 2007; Aristei et al., 2011; Lin et al., 2022), and in long-before primed conditions, relative to unprimed conditions (e.g., Huttenlocher & Kubicek, 1983; Sperber et al., 1979), may be due, entirely or partly, to easier and hence faster object recognition in related contexts. Although there are some indications in prior priming/blocking studies pointing to the perceptual notion, direct evidence for contextually facilitated object recognition has been lacking so far.

In this study, we have collected experimental data that can help understand what happens at the entry-level processing stage in blocked/primed picture-naming tasks. As overall naming times to pictures have the inherent shortcoming to be potentially influenced by a number of different processes, the present investigation examined a different indicator of the perceptual effort required for object recognition. Instead of measuring how fast pictures can be identified and named, we measured how much visual detail of the picture must be processed to name the depicted object correctly. According to the current understanding of visual perception, visual processing tends to go from coarse- to fine-grained information over time (for a review, see Hegdé, 2008). Thus, the less fine-grained visual information is required for object recognition, the earlier object recognition should be reached during picture processing. To detect the level of visual detail that suffices for object recognition, we simulated the temporal progression of gradual coarse- to fine-grained perception in a controlled manner. We presented all pictures to be named in gradually de-blurring image sequences. The de-blurring level that first allowed for correct object naming was measured as the recognition threshold of the object and served as an index to the relative effort and time required for recognition in the respective context condition.

In order to provide a comprehensive view of the results that we obtained and the insights they can provide, the General Discussion is split into two parts. We begin with a discussion of our main findings within the scope of prior empirical evidence and theoretical work from object recognition research. This discussion should help to understand how recognition is generally accomplished in blocked/primed picture-naming tasks and which factors influence the recognition process. In the second part of the General Discussion, we discuss how the discovered recognition effects may bear relevance to issues of word production and the interpretation of picture-naming latencies. We will summarize which aspects of perceptual accounts of semantic facilitation effects have received support from our data and which aspects remain to be tested in future research. In the last two subsections, we will sketch some broader implications of the detected recognition effects for the interpretation of semantic context effects obtained in the cyclic semantic blocking paradigm.

Visual Object Recognition in Blocking and Priming Paradigms

Object Recognition Succeeds on the Basis of a Partially Analyzed Version of the Picture

The results of both experiments have shown that objects can be successfully identified on the basis of relatively coarse-grained visual information, in all contexts and irrespective of the category type. This indicates that a full visual analysis of all pictured information is generally not required in word production tasks in which objects have to be named on a rather unspecific (often called: basic) level.

The finding that basic-level objects can be identified without access to and processing of fine-grained picture detail does not seem to be a paradigm- or task-specific effect. The finding is consistent with the current evidence on visual object perception. Studies on peripheral vision have demonstrated that basic level categorizations do not require direct foveation of the target but can succeed even in the visual periphery where the spatial acuity is reduced, similar to what our blurring levels have simulated. Ramezani et al. (2019), for example, found that participants were able to categorize objects on a “pigeon” or “racer” level even at 18° of eccentricity with an accuracy of about 75%.

Current theories of visual perception consider such findings as evidence for the visual percept being a product of incoming sensory information and knowledge stored in memory (e.g., Bar, 2007, 2009; O’Callaghan et al., 2017). Any incoming sensory information seems to be continuously proactively linked to analogous representations stored in memory. This mechanism triggers the activation of “perceptual hypotheses” (Trapp & Bar, 2015), that is, object representations that are considered as most likely interpretations of the given image, whereby higher-level information of expected object representations becomes activated before the respective information is bottom-up accumulated (Bar et al., 2006). The pre-activation of expected object representations substantially facilitates the formation of a meaningful percept. The visual system can test the top-down pre-sensitized object representations against incoming information and as soon as one of these candidate representations matches well with incoming information, the respective perceptual hypothesis seems to be confirmed. The classic text book example used to demonstrate such top-down effects is Gregory's (1970) "Hidden Dalmatian" picture. Although the picture just shows a black-and-white spotted pattern, people recognize a dog in the picture, at the latest with the help of the title "Dalmatian." The expectation to likely see a dog activates stored knowledge about dog shapes. This higher-level information matches well with incoming lower-level information from bottom-up streams (a particular pattern of spots) and lets people perceive a dog although no dog-specific features are shown in the picture. Sinha and Adelson (1997) demonstrated that such top-down effects can be extremely strong and even over-facilitate recognition when the set of perceptual hypotheses includes more or less just one candidate interpretation of the input. They manipulated the Dalmatian picture by modifying black spots critical for a dog shape and, although their manipulated picture no longer supported a dog interpretation, some participants nevertheless reported to have seen the dog. This finding indicates that in trying to make a stimulus fit a strong expectation (here: to see a dog), people can apparently become blind to ambiguous or conflicting information, making them mistakenly believe the pictured information was compatible with the interpretation they anticipated.

The crucial implication of the finding that object recognition generally succeeds on the basis of a partially analyzed version of the picture is that object recognition in picture-naming tasks may always enjoy some top-down facilitation from a certain point in time, just due to the predictive nature of the visual system. This enables participants to reduce the relatively effortful bottom-up analysis to a level that is just sufficient to identify the object.

The Superordinate Category Is Recognized Before the Object

The results of our explorative analysis in Experiment 1 have shown that naming responses converged on the target's semantic category relatively early and only thereafter, based on more precise visual input, participants were able to identify and name the concrete object. This pattern was found in all contexts and irrespective of the category type, suggesting that recognition may generally start with some broad taxonomic classification, which then becomes progressively more specific up to the level that is needed for task demands.

Our category-before-object finding is in line with recent findings on object perception. A large and growing body of studies (e.g., Clarke et al., 2013, 2015; Gauthier et al., 1997; Grill-Spector & Kanwisher, 2005; Poncet & Fabre-Thorpe, 2014; Ramezani et al., 2019; Thorpe et al., 1996; VanRullen & Thorpe, 2001) has demonstrated that broad categories of objects (e.g., knowing that an object is an animal) can be recognized at a glance, that is, ultra-rapidly, and solely on the basis of coarse information, for example, also in far peripheral vision. Identifying the object on finer taxonomic levels (e.g., knowing that it is a bird or even a pigeon), however, requires vision with, respectively, longer/more scrutiny and accessibility to more fine-grained object information. Such evidence has supported the view that in perceiving a stimulus in increasingly greater detail over time, we place it in categories of increasing specificity (cf. also the MEG decoding results of Ritchie et al., 2015, demonstrating that the decoding accuracy increased with more processing time, indicating a gradually improved neural separability of categorically related objects over time).

In top-down frameworks, initial categorizations are seen as an important trigger of top-down streams that aid the subsequent recognition (see e.g., Bar, 2003; Fenske et al., 2006; Hochstein & Ahissar, 2002; Schendan & Ganis, 2015; Trapp & Bar, 2015). Arriving at an "initial guess" as to the object's category membership allows the visual system to anticipate just a minimal set of likely objects. These likely object representations get anticipatorily enhanced and object representations from irrelevant categories, even those with global profiles similar to the input, can get top-down suppressed (cf. Trapp & Bar, 2015). This mechanism substantially limits the number of perceptual hypotheses that need to be considered when further visual signals come in: just a few most likely candidates from a single category need to be tested. This reduces the total processing effort and simplifies the recognition process to a within-category discrimination task. The observer just has to decide which category member may be the relatively best interpretation of the input.

Objects From Visually Similar Categories Are Comparatively Difficult to Recognize

In both experiments, we obtained a significant main effect of the visual type of the semantic category, indicating that intra-category visual similarity has a general, negative impact on object recognition. Objects from visually similar (VS) categories always required a more extensive and fine-grained picture processing than objects from visually dissimilar (VD) categories did.

Effects of intra-category visual similarity on perceptual processes were also found in prior studies (e.g., Alexander & Zelinsky, 2011; Ashworth III & Dror, 2000; Humphreys et al., 1988; Laws & Gale, 2002). A theoretical explanation of the relative disadvantage of intra-category visual similarity can go as follows. After a first broad categorization of the stimulus is achieved, the visual system can constrain the space of perceptual hypotheses to a minimal set of category exemplars from a single category. The ease with which the hypothesis space can be further narrowed thereafter may however differ, depending on the category. While salient and relatively rapidly perceived global information suffices to discriminate candidate representations of VD categories (e.g., a lighthouse from a pyramid or windmill), this type of information is not yet sufficient for object discriminations within VS categories (e.g., to discriminate an eagle from an owl or parrot which have similar global structures). To exclude inaccurate candidates of VS categories, the observer is required to further analyze the picture. In addition to the global information also some of the less salient finer picture details need to be analyzed, because only these details provide distinctive information. The detection of local information is, however, relatively slow (e.g., Navon, 1977; Schyns & Oliva, 1994) because focused attention is needed to bring details, their location, and precise conjunction into consciousness (Hochstein & Ahissar, 2002; see also Treisman, 1988). The fact that distinctive features are small and few for VS objects, but common features are numerous and salient, could make the detection of critical details even more difficult (cf. set-size effects observed in visual search tasks, e.g., Chong & Treisman, 2005), because salient information may distract attention

(J. M. Wolfe & Horowitz, 2017), which makes the top-down guidance of attention (Hochstein & Ahissar, 2002) less successful. Last but not least, the decision whether incoming information matches a candidate's representation sufficiently well or still too insufficiently may be more difficult when small deviations matter. While a rough estimate of match accuracy may be sufficient in the case of VD categories, comparatively precise computations might be required when candidates of VS categories need to be discriminated (cf. the higher difficulty of distractor rejections in visual search tasks when the contrast between target and distractors is low, J. M. Wolfe & Horowitz, 2017).

The detected category type effect is potentially relevant for picture-naming studies in that it signals that the ease of object recognition is not only a matter of the presented stimulus. Instead, the detected category type effect suggests that both the presented stimulus and our perceptual knowledge about categories associated with the depicted object can influence how quickly participants know which object should be named.

Semantically Blocked and Primed Contexts Facilitate Object Recognition

The most important finding of the present study is that our data show a context-specific ease of object recognition in blocking and priming paradigms. We found that the amount of visual detail minimally sufficient for object recognition was significantly lower in primed/blocked contexts than in the respective unrelated contexts. This indicates that perceptual recognition processes work indeed more efficiently when the target is preceded by categorically related items.

The effect that blocking/priming contexts lower the perceptual recognition threshold of an object has never been documented before. In view of the observation that contextual facilitation emerged in both pictorial (Experiment 1) and non-pictorial contexts (Experiment 2), we consider a top-down explanation of the semantic context effect as most plausible and preferable to bottom-up visual priming accounts (for arguments against bottom-up priming, see the individual Discussion sections). The observation that contextual facilitation was the largest in our priming experiments further supports this view as this finding indicates that predictability matters and likely gives rise to the facilitations.

Enhanced top-down facilitation due to contextually improved predictability of the object to be recognized is in principle an established finding in the perceptual literature and an effect that has been observed in different paradigms. As some of these effects have striking parallels to the blocking/priming effect that we have found, we suggest to explain the detected blocking/priming effect along the same lines: through top-down mechanisms associated with a priori triggered expectations.

Findings from studies on visual search for instance have shown that prior cues where the target will be found (where-cues) or how the target will look (what-cues) facilitate the detection of basic features, feature conjunctions, or whole items (e.g., Gould et al., 2007; Newell et al., 2005; Vickery et al., 2005; J. M. Wolfe et al., 2004). Such findings suggest that a priori-triggered expectations improve the guidance of the visual attention (e.g., Hochstein & Ahissar, 2002; Trapp & Bar, 2015). Focused attention can immediately be directed to probably critical image regions. This makes already the earliest stages of the visual recognition process more effective and significantly constrains the bottom-up analysis that needs to be performed.

Early expectations about the target also seem to aid the disambiguation of the initially perceived coarse-grained visual signals. Bar (2004, see also Fenske et al., 2006) demonstrated that an impoverished low spatial frequency image (a blurred picture of an object) can be perceived as a hairdryer or a drill, depending on whether it appears in a bathroom or a workshop context. Bar and colleagues interpret this effect as a biasing effect that results when contextually triggered top-down expectations pre-exclude some possible interpretations of the input. A coarse input image that is actually ambiguous, that is, maps to multiple object representations such as a hairdryer, a drill, a pistol, etc., can become recognizable in particular contexts because of top-down expectations—here: predictions about which objects are likely to appear in the given setting—that further constrain the perceptual hypothesis spaces. Object representations that are considered unlikely given the current context are top-down suppressed. This suppression reduces the candidate space and processing is biased in favor of the most probable hypothesis, apparently even without accumulating further, more distinctive bottom-up information (see also Trapp & Bar, 2015). Similar biasing effects of contextually triggered expectations have also been reported by Bar and Ullman (1996). They presented ambiguous object drawings and observed that the initially ambiguous objects were recognized when shown together with stereotypical spatial context information.

Studies on peripheral vision have demonstrated that a priori-triggered expectations can also have a third benefit. They seem to help produce information-rich signals. B. A. Wolfe and Whitney (2014) and Harrison et al. (2013) investigated recognition performance in peripheral vision and found that prior cues to peripheral, crowded target objects improved their identification, even when a direct foveation, that is, clear and sharp perception, of the target was prevented. Harrison et al. (2013) suggested to explain this effect by assuming that the visual system can effectively "presample" an object that is expected to be seen soon (in foveal vision), a mechanism that Herwig and Schneider (2014) described as a transsaccadic feature prediction that builds on previous input and stored object knowledge. Top-down, and especially predictive coding, frameworks for visual perception (see e.g., Clark, 2013; O'Callaghan et al., 2017; Trapp & Bar, 2015) consider such effects to be a general mechanism of the visual processing system for the purpose to accelerate identification. The visual system seems to constantly attempt to refine the visual percept by testing the most likely perceptual hypotheses against incoming lower-level representations of the input (Trapp & Bar, 2015). Any higher-level information that goes well with the bottom-up information can potentially increase the meaningfulness of the resultant visual percept, which helps to recognize the target faster.

We assume that all just described top-down mechanisms will also arise in primed/blocked contexts, adding up to the relatively higher recognition efficiency in related contexts. It appears that the prime stimulus or the preceding item(s) in homogeneous blocks have an effect like prior cues; they allow the visual system to generate predictions about the target in advance so that already the earliest stages of the perceptual process (operating on coarse visual information) become top-down improved. The predictions initiating top-down streams may be generated as follows: Primed/blocked contexts give a cue to the category of the upcoming target. This triggers, guided by taxonomic associations, top-down expectations about which objects are likely to appear in the upcoming picture (typical exemplars of the expected category) and which are not. According to these expectations, a minimal set of object representations may get anticipatorily pre-sensitized, all others top-down suppressed, and picture processing can start with a minimal space of perceptual hypotheses related to just one category. By contrast, picture processing in unrelated contexts has to start without any top-down facilitation. As unrelated contexts do not provide any predictive information in advance, no candidate representations can be pre-sensitized and initially incoming signals necessarily have to be compared with almost all object representations stored in memory. This increases the amount of computation (cf. Bar, 2003) and makes especially the early stages of the visual process less effective. A similarly small, category-constrained hypothesis space might be achieved in unrelated contexts only after the bottom-up analysis has yielded an initial categorization of the input.

We assume the a priori reduction of the hypothesis space in blocked/primed contexts to be key to the increased recognition efficiency in related contexts. Top-down signals can improve already the initial coarse perception of the input. This makes early low-level representations of the visual input more meaningful and a quite cursory analysis of the pictured information potentially sufficient for object recognition, implying that participants could potentially reduce their visual bottom-up analysis more strongly in related contexts than in unrelated contexts.

Factors That Further Enhance Top-Down Context Effects on Object Recognition

Our data show that the magnitude of context effects increases with two factors: intra-category visual dissimilarity and explicitness of the advance category information. VD categories benefited significantly more from categorically related contexts than VS categories did and semantic context effects were significantly larger when the context explicitly named the superordinate category of the target (superordinate word primes) than when it showed single exemplars from the target category and the participant had to deduce the categorical relation (homogeneous blocks). An additive relationship between the two factors was at least numerically observed. Context effects were the largest for VD categories in our priming experiment and the smallest for VS categories in our blocking experiment.

The beneficial role of intra-category visual dissimilarity seems to be a logical consequence of the general negative visual similarity effect on object recognition (see section Objects From Visually Similar Categories Are Comparatively Difficult to Recognize). Objects from VD categories differ in their global profiles. As global information is initially perceived, the benefit of a category-constrained hypothesis space already during the initial cursory analysis of the stimulus is particularly strong for these categories, leading to the situation that VD objects in related contexts require the least recognition effort overall.

The second observation that explicitness of a priori category information matters for facilitations is in line with results obtained by J. M. Wolfe et al. (2004, Experiment 6) in a visual search experiment. Wolfe and colleagues presented subjects with search displays with three to nine photographs of real objects and observed that object search, for example, for a cherry was more facilitated by prior category word cues (e.g., "fruit") than it was facilitated by prior picture cues showing another category exemplar (e.g., an apple). Wolfe and colleagues attributed this effect to the higher level of certainty of the predictive information in the case of category labels. Knowing that the target will be a fruit helps more that knowing that the target could perhaps be a fruit or another edible thing. This idea is transferable to priming and blocking paradigms. After processing the superordinate word prime, participants could be quite sure from which category the target would be, especially in our paradigm where the prime was reliably related to the target. In homogeneous blocks, on the other hand, category predictions might have been vaguer since single objects seen block-initially can give rise to various category associations.

In the top-down framework, certainty is considered to be a factor influencing how strongly a hypothesis space can be reduced (see, e.g., Bar, 2003), thus how effective top-down signals can enhance recognition. If predictive information is less reliable or imprecise, as in homogeneous blocks, broad and less accurate hypotheses cannot be a priori excluded, but must be initially considered as possible interpretations of the input. Hence, the a priori category-constrained hypothesis space might have been slightly larger in homogeneous blocks than in primed contexts, with the consequence that initial top-down effects might have been slightly less effective and recognition hence a bit more effortful in the blocking paradigm as compared to the priming paradigm (cf. Hick's law, 1952: a larger number of choices increases decision times).

In the remainder of the General Discussion, we will discuss the implications of our findings on object recognition for picture-naming latencies. We start by discussing how previously proposed perceptual accounts of semantic facilitation can be updated.

Implications for Perceptual Accounts of Semantic Facilitation Effects

Our findings confirm a number of assumptions made in perceptual accounts. First and foremost, it can be confirmed that context manipulations by semantic blocking and priming indeed affect the ease of object recognition. As hypothesized by perceptual accounts (e.g., Abdel Rahman & Melinger, 2007; Aristei et al., 2011; Huttenlocher & Kubicek, 1983; Lin et al., 2022; Sperber et al., 1979), we found facilitated object recognition in categorically related contexts relative to unrelated contexts. The more specific hypothesis of Sperber et al. (1979) that contextually facilitated object recognition may result from comparatively lower evidence requirements in the recognition system is also supported by our results. The de-blurring threshold levels permitting object recognition in related versus unrelated contexts clearly demonstrated that a certain amount of bottom-up information that allowed for object recognition in related contexts was not yet sufficient to recognize the same object in unrelated contexts. In order to reliably identify an object in unrelated contexts, participants had to analyze a bit more fine-grained pictured information, indicating that the visual system apparently demands the accumulation of relatively more bottom-up evidence in unrelated contexts to reach the point of recognition.

As any additional amount of computation needs additional time (cf. Bar, 2003), especially when more of the slowly working "vision with scrutiny" is demanded (Hochstein & Ahissar, 2002), our data support the claim that context effects on picture-naming times might be related to recognition effects. The lower recognition thresholds of objects in related compared to unrelated contexts that we observed should give related contexts a comparative advantage in terms of recognition speed and, as recognition times are entirely included in picture-naming times (picture-naming times are measured as the time between the onset of picture presentation and voice onset), also in terms of overall response speed. Thus, all considered, our data provide supporting evidence for the main claim of perceptual accounts. Recognition effects will likely cause faster picture naming in related contexts relative to unrelated contexts. How large this time advantage due to easier recognition in related contexts is can, however, not be estimated based on our data since our experiments did not include any time measurement. This leaves the question open whether recognition effects are the driving force of semantic facilitation effects in latencies. Future research will have to establish how fast recognition is reached in the critical blocking/priming conditions, and whether the differences in recognition speed can fully account for the differences observed in overall naming latencies.

With regard to the mechanism responsible for contextually facilitated object recognition, our data corroborate top-down accounts (e.g., Abdel Rahman & Melinger, 2007; Aristei et al., 2011; Lin et al., 2022). Our category contrast (VD vs. VS categories) showed that visually similar object features do not boost but attenuate contextual facilitation and the contrast of pictorial (blocking) versus non-pictorial contexts (priming with category words) showed that contextual facilitation does not hinge on pictorial contexts, but occurs in both pictorial and non-pictorial contexts. Both findings speak against bottom-up enhanced object recognition in related contexts, but can readily be accounted for by top-down effects triggered by predictive context preceding the target.

Note, that although our data do not provide any evidence for bottom-up enhanced recognition, this does not exclude that under different experimental conditions, significant effects of bottom-up visual priming may occur. Assuming, for example, that visual priming effects are relatively short-lived, our paradigms would have disfavored visual priming effects as the categorically related object pictures in Experiment 1 and the category names in Experiment 2 were presented up to several seconds before the first recognizable target picture was encountered. In single-picture-naming tasks, the context-target interval is much shorter, giving short-lived priming effects potentially more scope to affect perceptual processes. Another possibility why our data do not provide any evidence for bottom-up priming could be that bottom-up priming effects primarily influence the speed of visual information processing, but only marginally what and how much visual information needs to be processed. As our paradigms have only measured the latter, it is conceivable that bottom-up priming effects actually occurred, even in our paradigms, but our dependent variable was just not sensitive to them. On the other hand, Sperber et al. (1979) hypothesized that in pictorial contexts—the only contexts that should allow for bottom-up visual priming—an additional lowering of evidence requirements is to be expected, an effect that we have definitely not found. We, therefore, think that bottom-up priming hardly plays a role in causing semantic facilitation effects. The main reason why semantic facilitation emerges seems to be top-down expectations triggered by predictive contexts before the target. We could however imagine that bottom-up priming effects may possibly have an additional modulating effect on response times when pictorial and non-pictorial contexts with the same explicitness of advance category information are contrasted.

A last point that earlier perceptual accounts touched on and which now can be substantiated concerns the subprocess(es) of object recognition that become(s) more efficient in related contexts. It has been proposed that an increased overall efficiency of object recognition could in principle result from a more efficient bottom-up analysis of pictures or from a facilitated conceptual interpretation/categorization of the pictured information, or maybe even both (see Huttenlocher & Kubicek, 1983; Lin et al., 2022; Sperber et al., 1979). Our data provide direct evidence for an enhanced conceptual interpretation of pictured information: we observed that participants interpreted low-level visual information significantly better in related than in unrelated contexts. An improvement in the bottom-up analysis of pictures in related contexts is also likely given our results pointing to top-down effects, though our data did not directly demonstrate it. For one thing, top-down streams are thought to affect the guidance of visual attention (e.g., Hochstein & Ahissar, 2002; Trapp & Bar, 2015) and there is no reason to assume that this should be different in our paradigm. Secondly, the interaction of semantic context and intra-category visual similarity that we have found demonstrates that top-down context effects decrease when the anticipation of object-diagnostic features and their location is difficult (i.e., when candidates of VS categories are expected). This indicates that the efficiency of bottom-up processes in detecting critical object features apparently depends on the type of shape candidates that are expected. Thus, purely conceptual effects resulting from a limited set of probable conceptual interpretations (which might have been of comparable size for our VS and VD categories) can be only half the story of contextual facilitations; the other half seems to be related to how much prior expectations are able to improve bottom-up picture processing. To clarify whether top-down expectations in related contexts really affect the visual analysis of pictures to be named, it may be worth investigating eye movements during blocked/primed picture naming. If significant differences especially in the very first fixations can be found between related and unrelated contexts, an impact on the bottom-up analysis may be concluded.

Sperber et al. (1979) and Huttenlocher and Kubicek (1983) suggested that facilitation may arise “relatively automatically” in related contexts, simply as “a by-product” of the participants’ having processed the context (Sperber et al., 1979, p. 339). This assumption cannot be directly confirmed by our data, but it is in line with the current understanding of the visual system working like a “prediction machine” (e.g., Bar, 2007, 2009; Clark, 2013; O’Callaghan et al., 2017). The visual system seems to constantly generate and refine predictions about the upcoming input in order to minimize the relatively effortful and slowly working bottom-up processes. Hence, once predictive information is available, the visual system will likely utilize it to facilitate top-down the ongoing/upcoming perceptual process. We, therefore, surmise that contextually triggered top-down facilitation may be an inherent part of the visual information processing involved in blocked/primed picture naming, implying that a particular, conscious response strategy (as suggested by Belke, 2017; Belke et al., 2017; though for a different processing stage) might not be necessary for top-down facilitations to occur in related contexts.

Are Recognition Effects the Key Determinant of the Effects Seen in the First Cycle of Cyclic Semantic Blocking Paradigms?

What is clear from our results is that there is a semantic context effect on perceptual recognition processes and this context effect will likely shorten recognition times for the object to be named in categorically related contexts relative to unrelated contexts. What is not clear solely on the basis of our results is whether recognition effects bear the main responsibility for semantic facilitation effects obtained in naming latencies in the critical primed/blocked conditions. As a first step toward finding an answer to this question, we discuss in the following whether key findings on early semantic facilitation in blocked-cyclic naming experiments, which have previously been seen as evidence for a semantic origin of facilitation, could in principle be the result of top-down effects on object recognition.

Semantic distance/similarity effects on early semantic facilitation in blocked-cyclic naming (e.g., Lin et al., 2022; Navarrete et al., 2012, Experiment 2) are often considered as strong arguments that facilitation effects in cycle 1 are due to semantic priming: the higher the number of semantic features that are shared between-category members, the greater the strength of mutual co-activation and the faster the picture naming (cf. Navarrete et al., 2012). Abdel Rahman and Melinger (2009, see also 2019) pointed out that there is also another property of close categories that matters for semantic similarity effects: Close categories (e.g., ungulates) consist of fewer category members than broad categories (e.g., animals). We think this feature of close versus distant categories might be relevant for perceptual processes because it determines how great the predictive potential of homogenous blocks is. The smaller size of close categories causes block-initial items in close conditions (e.g., horse, donkey—next target: sheep) to be more precise category cues to the target than are block-initial items in broad conditions (e.g., owl, ant—next target: sheep; item examples taken from Lin et al., 2022). This leads to the situation that the visual system can anticipate the upcoming input relatively better for close categories. The a priori-defined search spaces might be smaller and the pre-sensitized perceptual hypotheses more accurate for close categories, relative to broad ones, increasing the effectiveness of top-down streams for close categories and making object identification relatively faster. Thus, semantic distance effects on early semantic facilitation do not necessarily need to be semantically caused; the same pattern of results would also be expected when recognition effects are the main driving force of the effects seen in cycle 1.

The finding that the amount of facilitation in naming latencies increases with ordinal position within the first cycle of homogeneous blocks (Navarrete et al., 2014) can likewise be related to top-down recognition effects and does not need to be a reflection of cumulative semantic priming as proposed by Navarrete and colleagues. As outlined in the previous sections, target predictability is relatively low at the beginning of homogeneous blocks since the category that will be shared by all items of the block can only be roughly guessed based on the first single object. With every further object seen in homogeneous blocks, however, expectations become progressively more precise. This improves the anticipation of likely targets throughout the trials of the first cycle and makes the perceptual hypothesis space increasingly smaller. The concomitant increase in the effectiveness of perceptual top-down processes would cause recognition times to decrease with ordinal position, and in consequence picture naming times as well.

The observation that semantic facilitation effects disappear when homogeneous blocks are not strictly homogeneous (Damian & Als, 2005; Navarrete et al., 2012) was originally attributed to short-lived automatic semantic priming effects that do not survive interleaving unrelated filler trials. Belke et al. (2017) pointed out that the absence of facilitation effects can equally well be attributed to the lack of top-down effects in semi-homogeneous contexts. When there is no way of reliably predicting the category membership of the upcoming item, top-down pre-activations cannot be initiated and facilitating top-down effects fail to emerge. Belke and colleagues proposed this notion primarily with regard to strategic lexical-semantic pre-activations that do not take place. Lack of target predictability might however affect perceptual processes as well. The visual system might not be able to generate expectations in advance that help identify the forthcoming visual input. Picture processing in semi-homogeneous blocks thus has to start as a more or less pure bottom-up process, making object recognition similarly difficult as in completely unrelated contexts, with the result that recognition times may not differ much between context conditions. An absence of semantic facilitation effects in latencies would thus be predicted by a perceptual top-down account of early semantic facilitation as well.

A last observation worthy of discussion is the relatively large diversity in the size of early semantic facilitation effects reported in prior blocking studies, ranging from null effects to significant latency effects of up to about 50 ms. When semantic facilitation effects would arise from largely automatic semantic priming within a categorical cohort, one might expect that they should occur reliably (in fact, as reliably as semantic interference effects do in repetition cycles) and always with similar strength. Knowledge about categorical relations is stored in long-term memory, thus categorical knowledge structures should not vary across studies or cycles. The observation that early semantic context effects significantly vary in size and can sometimes even be absent might therefore be quite difficult to square with accounts assuming automatic spreading activation to underlie facilitation. By contrast, it is relatively straightforward to reconcile variable effect sizes seen in cycle 1 with a perceptual top-down account. First, the results of our experiments have shown that the strength of top-down context effects on recognition processes differs significantly for visually similar (VS) and visually dissimilar (VD) categories. Whether homogeneous blocks consist of VS or VD categories is practically never controlled in word production studies and hence is a factor that likely varies within and between studies, giving rise to some variability in early semantic facilitation. Second, Belke (2017) observed in a comparison of blocking studies that facilitation effects in cycle 1 tend to be significant in studies with blocked context conditions, but insignificant in studies presenting homogeneous and heterogeneous blocks in random or alternating order. As already argued by Belke, this design effect on semantic facilitation is problematic for spreading activation accounts. In contrast, for top-down effects on perceptual processes, the order in which homogeneous and heterogeneous blocks are administered can be a relevant factor, because it affects the certainty about the type of the upcoming block. Blocked designs allow participants to reliably anticipate whether the upcoming block will be homogeneous or not, so it is clear whether the first trial of a block is predictive for the rest of the items or not. Conversely, mixed designs reduce the certainty about the type of block that is encountered. This decreases the certainty whether the initial item(s) of homogeneous blocks really provide helpful cues to the forthcoming targets. A-priori reductions of the perceptual hypothesis space may be considered as not sufficiently supported, with the consequence that object recognition benefits less, or not at all, from actually predictive contexts, and no significant context effect on recognition times, and finally response latencies, is obtained.

To conclude, it appears that several key findings on latency effects previously reported for cycle 1 of blocked-cyclic naming experiments could in principle result from modulations of the strength of top-down effects on object recognition. The perceptual proposal can capture the pattern of results and can explain why the emergence or absence, and the strength of semantic facilitation effects are susceptible to semantic and non-semantic factors. Moreover, as top-down facilitation of perceptual processes seems to be inherent in any object recognition process due to the predictive nature of the visual system, perceptual accounts seem to be the most parsimonious and straightforward explanation of latency effects in cycle 1. Additional (ad hoc) assumptions for certain conditions, for example, that participants could devise response strategies under certain design conditions, would no longer be necessary to account for the pattern of early semantic facilitation effects. This does not exclude that strategic or automatic semantic priming effects can also occur and will possibly contribute to the effects seen in cycle 1. However, given that facilitated object recognition in the first cycle of homogeneous blocks is clearly evidenced by our findings, whereas conclusive evidence for significantly facilitated lexical-semantic access is lacking at present (all arguments are based on overall naming latencies), it seems appropriate to favor the view that early semantic facilitation effects in cyclic blocking paradigms may primarily reflect perceptual recognition effects. If this view is correct, one might expect that absolute latencies and context effects in cycle 1 will be systematically and primarily modulated by variables influencing the recognition difficulty and the efficiency of top-down effects on recognition processes. Future research will have to clarify whether this is the case.

One factor influencing the difficulty of object recognition might be the type of picture stimuli, that is, whether objects are presented as line drawings, black-and-white photographs, or color photographs with neutral or natural scene background. Photographs provide more surface detail or even color information and both features have been shown to facilitate object recognition (Heuer, 2016; Price & Humphreys, 1989). Thus, it would be expected that absolute latencies in cycle 1 may vary as a function of the type of picture stimuli. The perceptual proposal would however not necessarily predict an effect of the type of stimuli on the size of the context effect (this is in line with the literature: previous blocking studies have obtained semantic facilitation effects of, for example, about 30 ms with color photographs (Abdel Rahman & Melinger, 2007), black-and-white photographs (Navarrete et al., 2014), and line drawings (Marful et al., 2014) alike). The reason is that the benefit of top-down effects may be smaller for photographic materials, relative to line drawings, because photographic materials allow for easier object recognition in the first place. Moreover, additional surface (and color) information seems to help distinguish visually similar objects (Price & Humphreys, 1989), thus availability of this type of information likewise potentially reduces the relative advantage of the top-down context effect. To clarify whether latencies and the size of facilitation effects in the first cycle are primarily determined by perceptual effects, it might thus be advisable to focus on alternative factors affecting the general ease of recognition and the effectiveness of perceptual top-down effects in the same way, such as intra-category visual similarity which, according to our results, seems to hamper both the general recognition and the effectiveness of top-down processes in the perceptual system.

Side-Effects of Top-Down Enhanced Object Recognition: May Semantic Interference Effects in Repetition Cycles of Blocking Paradigms Be Boosted by Recognition Effects?

In the previous sections, we have argued that contextually facilitated object recognition will likely play a key role in the pattern of latencies in the first cycle of blocked cyclic naming studies because easier object recognition shortens recognition times, which in turn shortens overall response times in homogeneous blocks. In within-block repetition cycles, recognition times might be of minor importance for latency effects. Repetition priming effects likely allow for easier recognition in general. This substantially shortens recognition times in repetition cycles compared to the first one (cf. the results reported by Lin et al., 2022: larger N1 amplitudes predict shorter latencies) so that any contextual influence on recognition times might be of little consequence for overall response latencies.

The loss of influence of recognition times however does not necessarily imply that recognition effects play no role at all in within-block repetition cycles. A possibility that remains is that recognition effects may have some indirect impact on latencies, namely by modulating the strength of concept activations and their distribution. Cascade models of object naming (Humphrey et al., 1988, 1995; Humphreys & Forde, 2001) propose that perceptual object activations are passed on to the semantic processing system. This suggests that the pattern of concept (co-)activations depends initially primarily on which and how much depicted information has been analyzed, encoded, and activated during the recognition process (see also Belke, 2013, making a similar assumption). Accordingly, relatively easier or more difficult recognition could in principle matter for the ease of lexical retrieval. Assuming that semantic interference effects in within-block repetition cycles are due to lexical competition, recognition effects could indirectly contribute to semantic interference if the perceptual activation of object representations would lead to a more competitive pattern of concept activations. Given our results from the blocking experiment, we can conceive of at least two effects associated with the recognition process that could potentially boost lexical competition in homogeneous blocks.

First, our finding that a quite cursory analysis of pictures suffices for object recognition in homogeneous blocks, but not in heterogeneous blocks, suggests that the number of visually encoded object features may be smaller in homogeneous blocks as compared to heterogeneous blocks. The higher recognition efficiency in homogeneous blocks might cause the perceptual target activation accumulated up to the point of recognition to be comparatively low, leading to a comparatively low target activation at the lexical level as well, so that the relative difference between the activation level of the target and that of competing lexical candidates will be relatively small, thus competition high. In heterogeneous blocks, by contrast, participants are forced to analyze the pictures in more detail. In order to reach the point of recognition, participants have to visually encode a relatively higher number of object features. This might increase the perceptual and in consequence the lexical-semantic activation of the target, relative to that in homogeneous blocks, and critically, in a context that suffers least from lexical competition effects anyway. Lexical retrieval would thus be relatively easier in heterogeneous blocks.

The theoretical consideration that differences in the perceptual demands of object recognition may support a pro-competitive activation pattern in homogeneous blocks on the one hand (due to a minimal, just sufficient target activation resulting from contextually facilitated object recognition), but an anti-competitive activation pattern in heterogeneous blocks on the other hand (due to a quite high target activation resulting from the more detailed picture analysis), is consistent with findings recently reported by Lin et al. (2022). Lin and colleagues observed that stronger P2/N2 responses (associated with more difficult lexical-semantic access) generally predict longer latencies in repetition cycles, but in homogeneous blocks of close categories this P2/N2 modulation on latencies was additionally modulated by the strength of the N1 response: P2/N2 modulations on naming latencies were most salient when N1 was strong, thus object recognition particularly easy. This finding supports our idea that effects related to recognition processes do not adversely affect lexical retrieval in heterogeneous blocks, but in homogeneous blocks contextually facilitated object recognition does hinder lexical retrieval, presumably because it exacerbates lexical competition.

The second way in which recognition effects potentially contribute to a pro-competitive activation pattern in homogeneous blocks is that they may support high activation levels of semantically related object representations, through top-down pre-activations (associated with the a priori-reduced hypothesis space in homogeneous blocks) and perceptual co-activation (due to intra-categorically shared object features). As described in section Semantically Blocked and Primed Contexts Facilitate Object Recognition, prior category expectations in homogeneous blocks trigger that the recognition process will likely start with a reduced space of perceptual hypotheses. This might help enhance the activation level in the target's semantic cohort in homogeneous blocks in two ways: first, representations of likely category exemplars get anticipatorily activated to some degree, independent from the actual input pattern, and second, only representations from the target category can receive some boost from bottom-up streams because only these representations are considered in matching procedures. In heterogeneous blocks, by contrast, initially, incoming signals have to be compared with almost all object representations stored in memory due to the lack of target predictability. Thus, all object representations of whatever category that share the global profile of the target will initially receive some activation boost. This makes the pattern of perceptually induced co-activation of non-target representations generally more broadly distributed in heterogeneous blocks as compared to homogeneous blocks, and the lack of any a priori pre-activation additionally helps keep the activation level within the target's semantic cohort low.

Perceptually induced co-activation of non-target representations, which is in principle an inherent part of every recognition process due to the initial ambiguity of incoming visual signals, might moreover be of particular consequence in homogeneous blocks if the members of the block-defining category have many or some especially salient visual features in common. In this case, the amount of ambiguous information that will be encoded until object recognition is reached might be quite high (cf. the visual similarity effects on recognition thresholds we have found), with every encoded visual feature that is common to several category exemplars increasing the activation of the target but also that of the respective alternative category members. The consequence will be that the total level of activation accumulated within the target's semantic cohort is high for perceptual reasons alone. Automatic activation spreading at the semantic level might further increase the activation level since mutually inter-related concepts enhance each other's activation (Abdel Rahman & Melinger, 2009, 2019), with the result being that the activation level of alternative category candidates might be very close to that of the target at the lexical level and lexical competition particularly high. A similar scenario of the effect of intra-category visual similarity has been sketched by Humphreys et al. (1995) based on the finding that visual similarity appears to affect response times more strongly in naming than in object decision/semantic classification. They proposed that the costs of visual similarity between-category members increase when activation is passed through increasingly more processing levels because effects from the early perceptual stage are passed on and added to effects at later stages.

Regarding the issue that intra-category visual similarity possibly boosts lexical competition effects, findings of blocking studies manipulating the semantic distance/similarity seem to provide interesting data. When semantic distance is manipulated, object features such as "has four legs" are usually seen as semantic object features which can distinguish semantically close categories whose members have this feature in common (e.g., ungulates: horse, donkey, sheep, camel, roe deer) from semantically broad categories whose members do not have this feature in common (e.g., animals: horse, butterfly, parrot, trout, gorilla; category examples taken from Lin et al., 2022). However, semantic features á la "has for legs," "has a beak," or "has a handle" are not only semantic object features, but also visual object features that increase the visual similarity within close categories as compared to broad ones. Indeed, some of our visually similar categories (e.g., upper body clothing, birds, ungulates, hand tools) have been employed as semantically close categories in previous studies (see e.g., Lin et al., 2022), giving rise to the question whether the larger semantic interference effects obtained for close categories, relative to broad categories (e.g., Lin et al., 2022; Vigliocco et al., 2002), really originated from effects at the semantic stage or rather from effects related to the recognition process (for a similar claim, see Vitkovitch et al., 1993). In our view, it seems likely that perceptual effects related to intra-category visual similarity contribute to the enhanced interference effects for close categories. Thus, in contrast to Damian et al. (2001) or Belke et al. (2005), we assume that visual similarity does matter for semantic interference effects. The reason is, however, not a visual confusion effect, as they discussed and ruled out, but a less effective recognition process in the case of visually similar categories which causes stronger perceptual co-activation within the target's semantic cohort, making target-related representations more potent competitors at the lexical level than they would be for semantic reasons alone.

In sum, assuming that the pattern of perceptual object (co-)activations is transmitted to the lexical-semantic system, it appears conceivable that effects associated with the recognition process bear some relevance to inhibitory context effects in within-block repetition cycles. Future research is, however, needed before a firmer conclusion can be reached regarding the role of perceptually induced concept (co-)activation in lexical competition effects.

Conclusion

The findings of the present study have shown that one of the techniques most commonly used to study lexical retrieval during word production, manipulation of the semantic naming context, does not only affect lexical-semantic processes in a significant way. Manipulations of the semantic context of picture naming can affect the perceptual effort required for object recognition as well. The results of the reported experiments demonstrate that object recognition is significantly facilitated when the target picture follows item(s) from the same semantic category, with the strength of facilitation increasing the more specific/reliable the prior contextual cues at the target object are and the lower the visual similarity within the target's semantic category is. As the occurrence of contextual facilitation of object recognition is independent of the modality of the context items (pictorial or non-pictorial), but the strength of facilitation is modulated by the degree of target predictability, the enhanced object recognition in related contexts is likely due to top-down effects based on prior expectations.

The notion that context manipulations by semantic blocking and priming could give rise to recognition effects is by no means a novelty (see Abdel Rahman & Melinger, 2007; Aristei et al., 2011; Huttenlocher & Kubicek, 1983; Lin et al., 2022; Sperber et al., 1979), but little has been done so far to directly test perceptual accounts of context effects. Our data on the perceptual recognition threshold of objects in related versus unrelated contexts provide direct evidence that recognition efficiency is higher in homogeneous blocks, relative to heterogeneous blocks, and in categorically primed conditions, relative to unprimed conditions. Accordingly, the notion of perceptual accounts that shorter naming latencies in the respective blocking/priming conditions may be due to recognition effects receives support. Moreover, we have argued that several key findings on semantic facilitation effects reported in previous blocked-cyclic naming studies seem to be attributable to top-down recognition effects, suggesting that perceptual recognition effects may be the main determinant of the pattern of naming latencies seen in cycle 1 of blocked-cyclic naming studies. If this view is correct, latency effects obtained in cycle 1 would not directly speak to issues of lexical-semantic access in word production. However, we wish to stress that the fact as such that homogeneous blocks facilitate object recognition, relative to heterogenous blocks, might bear relevance to theories of lexical-semantic access anyway, because differences in the efficiency of object recognition could possibly affect concept activations up to the point where lexical-semantic processes take over, and in doing so, they could possibly influence the ease of conceptual and lexical access in word production.

The challenge for future research is to examine the theoretical and methodological ramifications of the finding that object recognition in blocked/primed picture-naming tasks depends on the context and the target's semantic category. Given the crucial role that is attributed to context effects (and their modulations by semantic distance) in discussions of theories of lexical access, our findings raise two questions in particular. The first concerns the theoretical understanding of the interplay of perceptual and lexical-semantic processes in word production. Most of the current lexical access models consider the recognition system as a discrete processing system which does not directly affect lexical-semantic processes. However, if it turns out that perceptual recognition effects would in part be responsible for latency effects, it may be more appropriate to extend lexical access models to the recognition stage so that effects associated with the recognition process can be considered as a regular determinant of mechanisms operating at the semantic and lexical level.

The second question that arises from our findings is which consequences should follow from the observation that classic word production paradigms such as semantic blocking manipulate target predictability, and apparently give rise to predictability effects. The role of predictions and predictability in cognitive processes is currently a research topic of high interest, but we are only just beginning to understand which findings and phenomena are influenced by predictability, and how. Word production research has turned its attention to this topic relatively late. It thus remains a challenge for future work to examine to what extent processes involved in word production are influenced by predictions, and how these influences could be theoretically modeled.

References

Abdel Rahman, R., & Melinger, A. (2007). When bees hamper the production of honey: Lexical interference from associates in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 604–614. https://doi.org/10.1037/0278-7393.33.3.604

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes*, *24*(5), 713–734. https://doi.org/10.1080/01690960802597250

Abdel Rahman, R., & Melinger, A. (2011). The dynamic microstructure of speech production: Semantic interference built on the fly. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 149–161. https://doi.org/10.1037/a0021208

Abdel Rahman, R., & Melinger, A. (2019). Semantic processing during language production: An update of the swinging lexical network. *Language, Cognition and Neuroscience*, *34*(9), 1176–1192. https://doi.org/10.1080/23273798.2019.1599970

Alario, F.-X., Segui, J., & Ferrand, L. (2000). Semantic and associative priming in picture naming. *The Quarterly Journal of Experimental Psychology: Section A*, *53*(3), 741–764. https://doi.org/10.1080/713755907

Alexander, R. G., & Zelinsky, G. J. (2011). Visual similarity effects in categorical search. *Journal of Vision*, *11*(8), Article 9. https://doi.org/10.1167/11.8.9

Aristei, S., Melinger, A., & Abdel Rahman, R. (2011). Electrophysiological chronometry of semantic context effects in language production. *Journal of Cognitive Neuroscience*, *23*(7), 1567–1586. https://doi.org/10.1162/jocn.2010.21474

Ashworth, A. R. S. III, & Dror, I. E. (2000). Object identification as a function of discriminability and learning presentations: The effect of stimulus similarity and canonical frame alignment on aircraft identification. *Journal of Experimental Psychology: Applied*, *6*(2), 148–157. https://doi.org/10.1037/1076-898X.6.2.148

Bajo, M. T. (1988). Semantic facilitation with pictures and words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(4), 579–589. https://doi.org/10.1037/0278-7393.14.4.579

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, *15*(4), 600–609. https://doi.org/10.1162/089892903321662976

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*(8), 617–629. https://doi.org/10.1038/nrn1476

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280–289. https://doi.org/10.1016/j.tics.2007.05.005

Bar, M. (2009). Predictions: A universal principle in the operation of the human brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1181–1182. https://doi.org/10.1098/rstb.2008.0321

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 449–454. https://doi.org/10.1073/pnas.0507062103

Bar, M., & Ullman, S. (1996). Spatial context in recognition. *Perception*, *25*(3), 343–352. https://doi.org/10.1068/p250343

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

Bechberger, L., & Scheibel, M. (2020). Analyzing psychological similarity spaces for shapes. In M. Alam, T. Braun, & B. Yun (Eds.), *Ontologies and concepts in mind and machine: 25th International conference on conceptual structures* (Vol. 12277, pp. 204–207). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-57855-8_16

Belke, E. (2013). Long-lasting inhibitory semantic context effects on object naming are necessarily conceptually mediated: Implications for models

of lexical-semantic encoding. *Journal of Memory and Language*, *69*(3), 228–256. https://doi.org/10.1016/j.jml.2013.05.008
Belke, E. (2017). The role of task-specific response strategies in blocked-cyclic naming. *Frontiers in Psychology*, *7*(3), Article 1955. https://doi.org/10.3389/fpsyg.2016.01955
Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *The Quarterly Journal of Experimental Psychology Section A*, *58*(4), 667–692. https://doi.org/10.1080/02724980443000142
Belke, E., Shao, Z., & Meyer, A. S. (2017). Strategic origins of early semantic facilitation in the blocked-cyclic naming paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(10), 1659–1668. https://doi.org/10.1037/xlm0000399
Bürki, A., Elbuy, S., Madec, S., & Vasishth, S. (2020). What did we learn from forty years of research on semantic interference? A Bayesian meta-analysis. *Journal of Memory and Language*, *114*, Article 104125. https://doi.org/10.1016/j.jml.2020.104125
Carr, T. H., McCauley, C., Sperber, R. D., & Parmelee, C. M. (1982). Words, pictures, and priming: On semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(6), 757–777. https://doi.org/10.1037/0096-1523.8.6.757
Chong, S. C., & Treisman, A. (2005). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, *67*(1), 1–13. https://doi.org/10.3758/BF03195009
Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477
Clarke, A., Devereux, B. J., Randall, B., & Tyler, L. K. (2015). Predicting the time course of individual objects with MEG. *Cerebral Cortex*, *25*(10), 3602–3612. https://doi.org/10.1093/cercor/bhu203
Clarke, A., Taylor, K. I., Devereux, B., Randall, B., & Tyler, L. K. (2013). From perception to conception: How meaningful objects are processed over time. *Cerebral Cortex*, *23*(1), 187–197. https://doi.org/10.1093/cercor/bhs002
Costa, A., Alario, F., & Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychonomic Bulletin & Review*, *12*(1), 125–131. https://doi.org/10.3758/BF03196357
Crowther, J. E., & Martin, R. C. (2014). Lexical selection in the semantically blocked cyclic naming task: The role of cognitive control and learning. *Frontiers in Human Neuroscience*, *8*, Article 9. https://doi.org/10.3389/fnhum.2014.00009
Damian, M. F., & Als, L. C. (2005). Long-lasting semantic context effects in the spoken production of object names. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1372–1384. https://doi.org/10.1037/0278-7393.31.6.1372
Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review*, *10*(1), 111–117. https://doi.org/10.3758/BF03196474
Damian, M. F., & Spalek, K. (2014). Processing different kinds of semantic relations in picture-word interference with non-masked and masked distractors. *Frontiers in Psychology*, *5*, Article 1183. https://doi.org/10.3389/fpsyg.2014.01183
Damian, M. F., Vigliocco, G., & Levelt, W. J. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*(3), B77–B86. https://doi.org/10.1016/S0010-0277(01)00135-4
Davies, D., Sperber, R. D., & McCauley, C. (1981). Intelligence-related differences in semantic processing speed. *Journal of Experimental Child Psychology*, *31*(3), 387–402. https://doi.org/10.1016/0022-0965(81)90025-4
Dell, G. S., Martin, N., & Schwartz, M. F. (2007). A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming. *Journal of Memory and Language*, *56*(4), 490–520. https://doi.org/10.1016/j.jml.2006.05.007
Esterman, M., & Yantis, S. (2010). Perceptual expectation evokes category-selective cortical activity. *Cerebral Cortex*, *20*(5), 1245–1253. https://doi.org/10.1093/cercor/bhp188
Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149
Fenske, M. J., Aminoff, E., Gronau, N., & Bar, M. (2006). Top-down facilitation of visual object recognition: Object-based and context-based contributions. *Progress in Brain Research*, *155*(Part B), 3–21. https://doi.org/10.1016/S0079-6123(06)55001-0
Flores d'Arcais, G. B., & Schreuder, R. (1987). Semantic activation during object naming. *Psychological Research*, *49*(2), 153–159. https://doi.org/10.1007/BF00308681
Gauthier, I., Anderson, A. W., Tarr, M. J., Skudlarski, P., & Gore, J. C. (1997). Levels of categorization in visual recognition studied using functional magnetic resonance imaging. *Current Biology*, *7*(9), 645–651. https://doi.org/10.1016/S0960-9822(06)00291-0
Glaser, W. R., & Düngelhoff, F. J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 640–654. https://doi.org/10.1037/0096-1523.10.5.640
Gould, I. C., Wolfgang, B. J., & Smith, P. L. (2007). Spatial uncertainty explains exogenous and endogenous attentional cuing effects in visual signal detection. *Journal of Vision*, *7*(13), Article 4. https://doi.org/10.1167/7.13.4
Gregory, R. L. (1970). *The intelligent eye*. McGraw-Hill.
Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science*, *16*(2), 152–160. https://doi.org/10.1111/j.0956-7976.2005.00796.x
Harrison, W. J., Retell, J. D., Remington, R. W., & Mattingley, J. B. (2013). Visual crowding at a distance during predictive remapping. *Current Biology*, *23*(9), 793–798. https://doi.org/10.1016/j.cub.2013.03.050
Hegdé, J. (2008). Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology*, *84*(4), 405–439. https://doi.org/10.1016/j.pneurobio.2007.09.001
Herwig, A., & Schneider, W. X. (2014). Predicting object features across saccades: Evidence from object recognition and visual search. *Journal of Experimental Psychology: General*, *143*(5), 1903–1922. https://doi.org/10.1037/a0036781
Heuer, S. (2016). The influence of image characteristics on image recognition: A comparison of photographs and line drawings. *Aphasiology*, *30*(8), 943–961. https://doi.org/10.1080/02687038.2015.1081138
Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*(1), 11–26. https://doi.org/10.1080/17470215208416600
Hirschfeld, G., Jansma, B., Bölte, J., & Zwitserlood, P. (2008). Interference and facilitation in overt speech production investigated with event-related potentials. *NeuroReport*, *19*(12), 1227–1230. https://doi.org/10.1097/WNR.0b013e328309ecd1
Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*(5), 791–804. https://doi.org/10.1016/S0896-6273(02)01091-7
Humphreys, G. W., & Forde, E. M. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. *Behavioral and Brain Sciences*, *24*(3), 453–476. https://doi.org/10.1017/S0140525X01004150
Humphreys, G. W., Lamote, C., & Lloyd-Jones, T. J. (1995). An interactive activation approach to object processing: Effects of structural similarity, name frequency, and task in normality and pathology. *Memory*, *3*(3–4), 535–586. https://doi.org/10.1080/09658219508253164
Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, *5*(1), 67–104. https://doi.org/10.1080/02643298808252927
Huttenlocher, J., & Kubicek, L. F. (1983). The source of relatedness effects on naming latency. *Journal of Experimental Psychology: Learning,*

Memory, and Cognition, *9*(3), 486–496. https://doi.org/10.1037/0278-7393.9.3.486

Janssen, N., Hernández-Cabrera, J. A., van der Meij, M., & Barber, H. A. (2015). Tracking the time course of competition during word production: Evidence for a post-retrieval mechanism of conflict resolution. *Cerebral Cortex*, *25*(9), 2960–2969. https://doi.org/10.1093/cercor/bhu092

Keefe, D. E., & Neely, J. H. (1990). Semantic priming in the pronunciation task: The role of prospective prime-generated expectancies. *Memory & Cognition*, *18*(3), 289–298. https://doi.org/10.3758/BF03213882

Kuipers, J. R., La Heij, W., & Costa, A. (2006). A further look at semantic context effects in language production: The role of response congruency. *Language and Cognitive Processes*, *21*(7–8), 892–919. https://doi.org/10.1080/016909600824211

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/JSS.V082.I13

La Heij, W., Heikoop, K. W., Akerboom, S., & Bloem, I. (2003). Picture naming in picture context: Semantic interference or semantic facilitation? *Psychology Science*, *45*(1), 49–62.

Laws, K. R., & Gale, T. M. (2002). Category-specific naming and the 'visual' characteristics of line drawn stimuli. *Cortex*, *38*(1), 7–21. https://doi.org/10.1016/S0010-9452(08)70635-X

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Lin, H. P., Kuhlen, A. K., Melinger, A., Aristei, S., & Abdel Rahman, R. (2022). Concurrent semantic priming and lexical interference for close semantic relations in blocked-cyclic picture naming: Electrophysiological signatures. *Psychophysiology*, *59*(4), Article e13990. https://doi.org/10.1111/psyp.13990

Lindsay, P. H., & Norman, D. A. (1977). *Human information processing: An introduction to psychology*. Academic Press.

Lupker, S. J. (1988). Picture naming: An investigation of the nature of categorical priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 444–455. https://doi.org/10.1037/0278-7393.14.3.444

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 503–535. https://doi.org/10.1037/0278-7393.33.3.503

Marful, A., Paolieri, D., & Bajo, M. T. (2014). Is naming faces different from naming objects? Semantic interference in a face-and object-naming task. *Memory & Cognition*, *42*(3), 525–537. https://doi.org/10.3758/s13421-013-0376-8

Meinzer, M., Yetim, Ö, McMahon, K., & de Zubicaray, G. (2016). Brain mechanisms of semantic interference in spoken word production: An anodal transcranial Direct Current Stimulation (atDCS) study. *Brain and Language*, *157–158*, 72–80. https://doi.org/10.1016/j.bandl.2016.04.003

Miller, J., & Bauer, D. W. (1981). Visual similarity and discrimination demands. *Journal of Experimental Psychology: General*, *110*(1), 39–55. https://doi.org/10.1037/0096-3445.110.1.39

Navarrete, E., Del Prato, P., & Mahon, B. Z. (2012). Factors determining semantic facilitation and interference in the cyclic naming paradigm. *Frontiers in Psychology*, *3*, Article 38. https://doi.org/10.3389/fpsyg.2012.00038

Navarrete, E., Del Prato, P., Peressotti, F., & Mahon, B. Z. (2014). Lexical selection is not by competition: Evidence from the blocked naming paradigm. *Journal of Memory and Language*, *76*, 253–272. https://doi.org/10.1016/j.jml.2014.05.003

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383. https://doi.org/10.1016/0010-0285(77)90012-3

Neely, J. H., & Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective-retrospective processing theory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 24, pp. 207–248). Academic Press.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1003–1019. https://doi.org/10.1037/0278-7393.15.6.1003

Newell, F. N., Woods, A. T., Mernagh, M., & Bülthoff, H. H. (2005). Visual, haptic and crossmodal recognition of scenes. *Experimental Brain Research*, *161*(2), 233–242. https://doi.org/10.1007/s00221-004-2067-y

O'Callaghan, C., Kveraga, K., Shine, J. M., Adams Jr R. B., & Bar, M. (2017). Predictions penetrate perception: Converging insights from brain, behaviour and disorder. *Consciousness and Cognition*, *47*, 63–74. https://doi.org/10.1016/j.concog.2016.05.003

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Poncet, M., & Fabre-Thorpe, M. (2014). Stimulus duration and diversity do not reverse the advantage for superordinate-level representations: The animal is seen before the bird. *European Journal of Neuroscience*, *39*(9), 1508–1516. https://doi.org/10.1111/ejn.12513

Price, C. J., & Humphreys, G. W. (1989). The effects of surface detail on object categorization and naming. *The Quarterly Journal of Experimental Psychology, Section A*, *41*(4), 797–827. https://doi.org/10.1080/14640748908402394

Python, G., Fargier, R., & Laganaro, M. (2018a). ERP Evidence of distinct processes underlying semantic facilitation and interference in word production. *Cortex*, *99*, 1–12. https://doi.org/10.1016/j.cortex.2017.09.008

Python, G., Fargier, R., & Laganaro, M. (2018b). When wine and apple both help the production of grapes: ERP evidence for post-lexical semantic facilitation in picture naming. *Frontiers in Human Neuroscience*, *12*, Article 136. https://doi.org/10.3389/fnhum.2018.00136

Ramezani, F., Kheradpisheh, S. R., Thorpe, S. J., & Ghodrati, M. (2019). Object categorization in visual periphery is modulated by delayed foveal noise. *Journal of Vision*, *19*(9), Article 1. https://doi.org/10.1167/19.9.1

Ritchie, J. B., Tovar, D. A., & Carlson, T. A., & Bethge, M. (2015). Emerging object representations in the visual system predict reaction times for categorization. *PLoS Computational Biology*, *11*(6), Article e1004316. https://doi.org/10.1371/journal.pcbi.1004316

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1–3), 107–142. https://doi.org/10.1016/0010-0277(92)90041-F

Schendan, H. E., & Ganis, G. (2015). Top-down modulation of visual processing and knowledge after 250 ms supports object constancy of category decisions. *Frontiers in Psychology*, *6*, Article 1289. https://doi.org/10.3389/fpsyg.2015.01289

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, *5*(4), 195–200. https://doi.org/10.1111/j.1467-9280.1994.tb00500.x

Sinha, P., & Adelson, E. (1997). Doggone Dalmatian!. *Perception*, *26*(5), 667–667. https://doi.org/10.1068/p260667

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture–word processing. *Journal of Experimental Psychology: General*, *109*(4), 373–392. https://doi.org/10.1037/0096-3445.109.4.373

Snodgrass, J. G., & McCullough, B. (1986). The role of visual similarity in picture categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 147–154. https://doi.org/10.1037/0278-7393.12.1.147

Sperber, R. D., Davies, D., Merrill, E. C., & McCauley, C. (1982). Cross-category differences in the processing of subordinate–superordinate relationships. *Child Development*, *53*(5), 1249–1253. https://doi.org/10.2307/1129013

Sperber, R. D., McCauley, C., Ragain, R. D., & Weil, C. M. (1979). Semantic priming effects on picture and word processing. *Memory & Cognition*, *7*(5), 339–345. https://doi.org/10.3758/BF03196937

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. https://doi.org/10.1038/381520a0

Trapp, S., & Bar, M. (2015). Prediction, context, and competition in visual recognition. *Annals of the New York Academy of Sciences*, *1339*(1), 190–198. https://doi.org/10.1111/nyas.12680

Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology Section A*, *40*(2), 201–237. https://doi.org/10.1080/02724988843000104

VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*, *13*(4), 454–461. https://doi.org/10.1162/08989290152001880

Vickery, T. J., King, L. W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, *5*(1), 81–92. https://doi.org/10.1167/5.1.8

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*(3), B61–B69. https://doi.org/10.1016/S0010-0277(02)00107-5

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*(4), 422–488. https://doi.org/10.1016/j.cogpsych.2003.09.001

Vitkovitch, M., Humphreys, G. W., & Lloyd-Jones, T. J. (1993). On naming a giraffe a zebra: Picture naming errors across different object categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(2), 243–259. https://doi.org/10.1037/0278-7393.19.2.243

Wheeldon, L. R., & Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitor. *Journal of Memory and Language*, *33*(3), 332–356. https://doi.org/10.1006/jmla.1994.1016

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, *8*(2), 227–233. https://doi.org/10.1016/S0959-4388(98)80144-X

Wolfe, B. A., & Whitney, D. (2014). Facilitating recognition of crowded faces with presaccadic attention. *Frontiers in Human Neuroscience*, *8*, Article 103. https://doi.org/10.3389/fnhum.2014.00103

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 1–8. https://doi.org/10.1038/s41562-017-0058

Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*, *44*(12), 1411–1426. https://doi.org/10.1016/j.visres.2003.11.024

Zhang, Q., Feng, C., Zhu, X., & Wang, C. (2016). Transforming semantic interference into facilitation in a picture–word interference task. *Applied Psycholinguistics*, *37*(5), 1025–1049. https://doi.org/10.1017/S014271641500034X

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168–171. https://doi.org/10.1111/1467-9280.00430

Appendix A

Table A1

The Experimental Stimuli Used in Experiments 1 and 2

Experimental items	Category name (prime word in Exp. 2)	Category type	Median visual similarity (IQR)	Median semantic similarity (IQR)
Zange (pliers), Säge (saw), Axt (axe), Schaufel (shovel), Hammer (hammer)	Werkzeug (tool)	VS	4 (0.75)	5 (0)
Bluse (blouse), Jacke (hoody), Uniform (uniform), Mantel (coat), Hemd (shirt)	Oberkörperbekleidung (upper body clothing)	VS	5 (0)	5 (0)
Spülmachine (dishwasher), Toaster (toaster), Fernseher (TV), Mikrowelle (microwave), Waschmaschine (washing machine)	Elektrogerät (electrical appliance)	VS	4 (1)	5 (0)
Taube (pigeon), Adler (eagle), Eule (owl), Storch (stork), Papagei (parrot)	Vogel (bird)	VS	5 (0)	5 (0)
Sonnenblume (sunflower,), Palme (palm), Rose (rose), Tulpe (tulip), Löwenzahn (dandelion)	Pflanze (plant)	VS	5 (0.75)	5 (0)
Kuh (cow), Hirsch (stag), Ziege (goat), Esel (donkey), Pferd (horse)	Huftier (ungulate)	VS	5 (0)	5 (0)
Leuchtturm (lighthouse), Iglu (igloo), Pyramide (pyramid), Windmühle (windmill), Kirche (church)	Gebäude (building)	VD	2 (1.5)	5 (1)
Salzstreuer (salt cellar), Teller (plate), Teekanne (teapot), Schüssel (bowl), Glas (glass)	Geschirr (tableware)	VD	2 (1.75)	4 (0.75)
Auto (car), Roller (scooter), Bus (bus), Fahrrad (bicycle), Traktor (tractor)	Straßenfahrzeug (road vehicle)	VD	3 (1.75)	5 (0)
Paprika (sweet pepper), Spargel (asparagus), Zwiebel (onion), Brokkoli (broccoli), Möhre (carrot)	Gemüse (vegetable)	VD	2.5 (2)	5 (0)
Libelle (dragonfly), Ameise (ant), Marienkäfer (lady beetle), Fliege (fly), Schmetterling (butterfly)	Insekt (insect)	VD	3 (0)	5 (0)
Zitrone (lemon), Banane (banana), Himbeere (rasberry), Kirsche (cherry), Apfel (apple)	Obst (fruit)	VD	2 (2.5)	5 (0)

Note. The German object names and category names of the stimuli are given in the table, with English translations in parentheses. The pictures cannot be reproduced in print due to license restrictions, but all pictures and image sequences will be provided on request (contact: scheibel@uni-duesseldorf.de). Category Types are coded as VS = visually similar and VD = visually dissimilar. Within-category similarity values, median and IQR = interquartile range, are given using the numerical coding: 1 (*dissimilar*), 5 (*similar*). Visual similarity values are based on the results from Pretest 4, semantic similarity values are based on the results from Pretest 3 of Experiment 1.

Received November 10, 2021
Revision received November 15, 2022
Accepted November 16, 2022 ■