Effects of Anchor Distractors on

Eye Movements in Action primed Visual Search

Keywords: Anchor Objects, Visual Search, Scene Grammar, Object function

Goethe-University Frankfurt

Faculty 05 - Psychology and Sport Sciences

Modul: PsyBSc10 ExPra

Lecturer: ___________

Winter term _______

Submission date: ___________

____

____

___________@stud.uni-frankfurt.de

Abstract

The current opinion in psychology proposes that a scene is hierarchically structured and consists of objects that follow certain rules, which we refer to as *scene grammar*. Besides, the concept of *object affordances* suggests an interaction between visual and motor processes. Here, we combined these two concepts to explore whether this proposed hierarchical structure is represented in our behavior. Therefore, we tested whether hierarchical object-to object relations in scenes are reflected during an action primed visual search task. 11 university students participated in our eye tracking experiment in which they had to search for objects on a circular display. We found that *anchor objects* (large objects that predict the location of other objects) which were highly related to a target object were significantly more distracting than *local objects* (small objects near anchor objects) which were unrelated to the target object. Furthermore, eye movement measures descriptively showed a trend suggesting the proposed hierarchical structure. However, this trend was not significant, possibly because our study was underpowered. Nevertheless, we conclude that a scene's structure plays an important role in cognitive processing of the environment. Hence, our study contributes to the research that focuses on what computations we need to exploit real-world regularities.

Table of Contents

Effects of Anchor Distractors on Eye Movements in Action primed Visual Search

Early research on *visual search* was based on experiments in which participants searched for isolated targets that were randomly presented on blank backgrounds and surrounded by distractors. This model of visual search is called *classic guided search* (Wolfe, 1994; Wolfe et al., 2011). In classic guided search participants looked for targets of a certain size, color and orientation. Thus, they looked for targets with a limited set of basic attributes (Võ & Wolfe, 2013; Wolfe et al., 2011). Over time, research, as well as the scenes and objects used, became more natural, because search is an important task in the real world (Wolfe et al., 2011). Võ et al. (2019) stated that the most important guiding factor within real-world scenes was meaning. This is consistent with *cognitive guidance theory*, which proposes that attention is directed to scene regions that are relevant in a given situation and semantically informative (Henderson et al., 2018; Võ et al., 2019). Other research showed that even when a scene was not entirely visible, object-to-object and object-to-scene relations were activated (Bar, 2004; Oliva & Torralba, 2007). However, it has not been investigated yet whether the degree of relatedness between objects influences the activation. Therefore, in the current study, we tested how different hierarchical object-to-object relations affect visual search behavior.

Objects in our environment are never randomly arranged (Biederman et al., 1982; Võ & Wolfe, 2013; Wolfe et al., 2011). Instead, they follow certain rules that we learned over the course of our lives. This set of rules is referred to as *scene grammar* and makes assumptions about *semantic* as well as *syntactic* aspects and therefore proposes *what* objects should be *where* in a scene (Biederman et al., 1982; Võ, 2021; Võ et al., 2019). Scene grammar, where objects are constrained in scenes, can be compared to the linguistic grammar, where words are constrained by sentences (Võ et al., 2019). It could help us efficiently understand scenes, recognize objects and guide goal directed actions (Võ, 2021). More recently, it has been proposed that this set of rules could be structured hierarchically (Võ et al., 2019). An exemplary structure of a bathroom scene is shown in Figure A1. Scenes are divided into different clusters which we call *phrases.* Each phrase consists of one *anchor object* and spatially related – *local* – objects. Furthermore, anchor objects are special in that they can predict where other objects are in a scene and play a key role in object search (Võ et al., 2019). While local objects are smaller and more movable (e.g., a toothbrush), anchor objects are bigger and rather stationary (e.g., a sink) (Draschkow & Võ, 2017; Turini & Võ, 2022). Thus, if you are looking for your toothbrush, it is efficient to look at the sink while excluding the other phrases (Võ et al., 2019).

The question arises whether knowledge about spatial relations exists only because we often see two objects together. This would mean that our object-to-object associations are present because of statistical learning (Turk-Browne, 2012). For example, we know that a toothbrush is usually found next to a sink because we see them together. However, it could also be that we have this knowledge because we use them together. This idea of an interaction between visual and motor processes goes back to Gibson's (1979) concept of *object affordances*. Instead of just looking at objects, we also interact with them (Castelhano & Witherspoon, 2016). For example, we put our toothbrush next to the sink (instead of far away), because then brushing our teeth is more efficient (Võ, 2021). Therefore, Võ (2021) proposed that phrases are formed based on objects' functions and everyday actions.

In the present study we combined the ideas of a scene's hierarchical structure and object affordances in order to take a deeper look at the function of phrases in visual search. For example, Boettcher et al. (2018) conducted eye tracking experiments to investigate the role of anchor objects and found that anchor objects can guide visual search and lead to faster response times (RT). Here, we did not investigate whether anchor objects facilitate search. Instead, we used them as distractors when searching for an object that is more or less semantically related to the distractor. This semantic relatedness between objects was already studied by Turini and Võ (2022) with pairwise similarity judgments. However, they did not investigate the effect of relatedness on behavior. Here, we conducted an eye tracking experiment in which participants searched for a target object in the presence of a manipulated distractor object as well as random other objects. The search trials that we later analyzed were circular display searches instead of scene searches to minimize influence of spatial guidance from a scene's spatial structure. Nevertheless, 30 % of the trials were real-world scene searches to make search more natural and to interrupt search routines. This idea has not been tested yet and stems from logical considerations. What we do know, however, is that a quick glance at a scene activates stored knowledge about the typical structure of a scene (Võ et al., 2019). We hoped that this knowledge would also transfer to the display searches. To make search even more natural and in accordance with the idea that phrases are formed based on objects' functions (Võ, 2021), we primed the action of the target object. Thus, we cued the target objects with their functions instead of names or pictures. This is consistent with Castelhano and Witherspoon (2016) who showed that visual search is more efficient when an object is described by its function than by a picture alone. This way, we could investigate whether the primed actions activate a structured network of representations that is similar to Võ et al.'s (2019) proposed hierarchy.

We hypothesized that the higher the relation between target and distractor object, the more distracting the distractor would be while searching for the target object. This would be reflected in higher RTs (time between the onset of the search display and the keypress response), number of fixations (how often the distractor was looked at) and dwell times (how long the distractor was looked at). That idea goes back to classic guided search and is in line with Wolfe et al. (2011) who stated that the ability to identify an object is limited by similar items nearby. The difference is that in classic guided search physical properties played a role, whereas in our study semantic properties did. We would see that Võ et al.'s (2019) proposed hierarchy is reflected in our behavior if anchor object from the same scene (regardless of the phrase) were more distracting than local objects from other scenes. Additionally, within the same scene, anchor object from the same phrase would be more distracting than anchor objects from another phrase. Likewise, anchor objects from the same scene and phrase would be more distracting than anchor objects from another scene. Moreover, anchor objects from the same scene but different phrase would be more distracting than anchor objects from another scene. Whereas we assumed there would be no difference between anchor objects from another scene and local objects from another scene, since all these objects are equally semantically related to the target object.

Method

In order to follow the open science approach, our study design and its analysis was preregistered with AsPredicted (https://aspredicted.org/6K3_PPP).

Participants

A total of 11 students from Goethe University Frankfurt (eight women, three men; 19 to 26 years old, *M* = 22.18 years) participated voluntarily in our study, either for course credit or without compensation. Due to time constraints, we were not able to find the number of participants we have preregistered. Prior to the actual experiment, we tested the participants for color blindness. None of them were color blind and all participants had normal or corrected-to-normal vision. Further, they were unfamiliar with the study hypotheses and stimulus material and gave written informed consent before participating in the study.

Materials

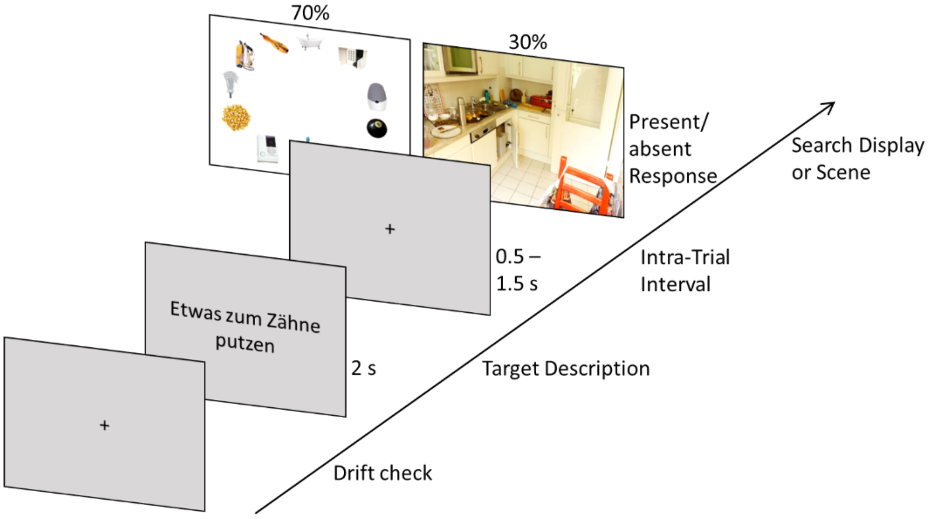
The stimulus material consisted of 1020 pictures of our target and anchor objects and was either taken from the Big and Small Objects database (Konkle & Oliva, 2012b) or from the web. Of each object, we used 10 typical and easy recognizable exemplars. Moreover, 2255 pictures of random distractor objects were collected from Brady et al. (2008) and through web search. To avoid confusion, there were no overlaps with the pictures of target and anchor objects. Furthermore, we did not use pictures of animals or human body parts because they could draw attention away from the other pictures. To avoid surrounding distraction, the backgrounds of all pictures were removed and set to white. Additionally, the pictures were resized to 130 x 130 pixels. Furthermore, we used 214 pictures of real-world indoor scenes from the BOiS database (Mohr et al., 2016a) which were resized to 1495 x 996 pixels. For each scene, there were two possible conditions: Either the target was located in a predictable place within the scene (knife on a plate), or the target was absent. Accordingly, we used 107 target-present scenes and 107 target-absent scenes and participants only ever searched in one condition of a scene. All pictures and scenes were presented in color. Furthermore, all search trials were randomized and counterbalanced for every four participants.

Procedure and Design

Participants performed an action primed visual search task on search displays or in real-world scenes. Prior to the actual experiment, they performed a few practice trials without eye tracking to familiarize themselves with the experiment. Afterwards, each participant completed the EyeLink calibration procedure. Eye positions were recorded while participants fixated a series of nine dots arranged in a square grid. The calibration was then validated using a second series of dots. After that, the actual experiment took place. An exemplary trial sequence is shown in Figure 1. Each trial began by looking at a fixation cross in the center of the screen while pressing the spacebar on the keyboard. This drift check ensured that there was no tracking error and initiated the next trial. Subsequently, the target object was described in German by its function. For example, a toothbrush was described as: "Etwas zum Zähne putzen" (in English: “Something to brush your teeth”). This description was

Figure 1

Exemplary Trial Sequence of the Experiment



Note. This is an exemplary trial sequence from the search phase of the experiment. Participants had to search for a target object either on a circular display or in scene and press "a" or "m" on the keyboard to indicate the presence or absence of the target object.

displayed on the screen for two seconds. Then, another fixation cross was displayed for 0.5 to 1.5 seconds before the search for the target object began. The search took place either on a search display (70 % of the trials) or in a real-world scene (30 % of the trials). The target object was either present (70 % of the trials) or absent (30 % of the trials). To indicate whether the target was present or absent, participants had to press "a" or "m" on the keyboard, respectively. Participants were instructed to respond as fast and as accurate as possible.

During the experiment, participants searched for 34 different target objects in 10 scene categories. Since no picture was shown more than once, we used different exemplars of the objects. This applied to both the target objects and the anchor objects. In each of the four conditions (related, unrelated, semantically unrelated and control), two different exemplars of a target object appeared. Thus, each target object appeared eight times. Furthermore, target and distractor positions were randomly assigned for each display configuration but held constant across conditions and participants. In order to make the search more difficult, the search display rotated mirrored. In this way, participants were not able to guess the position of the target object. For an individual participant, the experiment consisted of 272 trials and lasted about 60 minutes.

This study was designed as a within-subject design and was conducted in the Scene Grammar Lab at the Goethe University Frankfurt. We manipulated the relatedness between the target (e.g., a toothbrush) and the distractor objects. In accordance with Võ et al.'s (2019) proposed hierarchy we defined four conditions: Related, unrelated, semantically unrelated and control (see Figure A2). In the related condition, the anchor distractor (e.g., a sink) and the target object were from the same scene and phrase. In the unrelated condition, the anchor distractor (e.g., a bathtub) and the target object belonged to the same scene but to a different phrase. In the semantically unrelated condition, the anchor distractor and the target object were from a different scene. Whereas in the control condition the distractor was a random, non-anchor (local) object. We examined both behavioral (RT) and eye movement (number of fixations and dwell time) measures. However, for the sake of brevity, we deviated from the preregistered dependent variables.

Data Analysis

We analyzed data using R (R Core Team, 2022) in RStudio (Posit team, 2022). Trials with scene searches were excluded from the analysis since they were included only to stimulate the use of scene-based knowledge. In addition, we analyzed only correct trials in which the target object was present. RTs that were less than or more than three standard deviations above the mean were treated as outliers and discarded. After excluding the outliers, there were 1,759 trials (98.93 %) left for further analysis. The analysis of variance (ANOVA) with relatedness as factor was conducted using the *ez* package (Lawrence, 2016).

Since we wanted to compute a linear model, we visually checked whether our data met the assumptions. None of our dependent variables (RT, number of fixations and dwell time) were distributed normally. Thus, we log-transformed them before carrying out the ANOVA. We used the *ggplot2* package (Wickham, 2016) for graphics and the *emmeans* package (Lenth, 2022) for post-hoc comparisons.

Apparatus

Our experiment was conducted in a dimly lit room equipped with 22-in. monitors operating at a resolution of 1680 x 1050 pixels and a refresh rate of 60 Hz. A chin rest was used to maintain a viewing distance of 60 cm. Eye movements were recorded using the EyeLink 1000 Plus Desktop Mount (SR Research, Ontario, Canada) at 1000 Hz. Target, random distractor and anchor distractor interest areas were polygons containing the respective pictures on the display. Saccades and fixations were extracted from raw gaze data during recording by the EyeLink parser. Velocity and acceleration thresholds were set to the EyeLink default values of 30°/s and $8000°/s^2$ respectively. Further, we used PsychoPy (Peirce et al., 2019) for presentation of the experiment, stimulus processing and response recording.

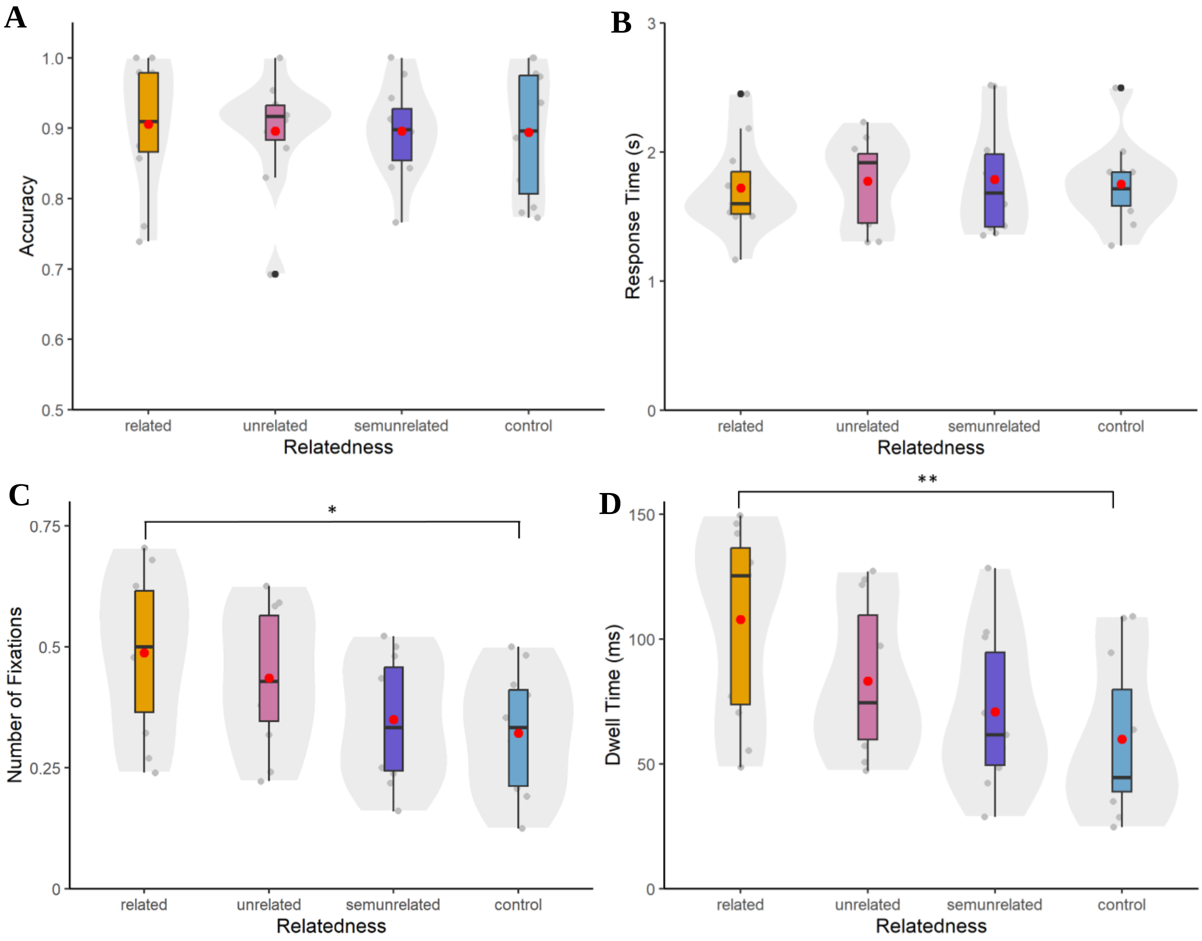
Results

Behavioral measures

As an explorative analysis, we examined the average search accuracy before excluding incorrect trials and found that it was quite high (M = 0.897, SD = 0.304) (see Figure 2A). Against our hypotheses, the ANOVA did not reveal a significant effect of relatedness on RT, $F(3,30) = 0.282$, $p = .838$, $\eta^2 = .005$. This indicates that participants were not distracted by an anchor object in the sense that their average RTs significantly differed between the different types of anchor distractors (see Figure 2B).

Figure 2

Effects of Relatedness on Accuracy (A), RT (B), Number of Fixations (C) and Dwell Time (D)



Note. The violin plots show the distribution of the data for Accuracy (A), RT (B) number of fixations (C) and dwell time (D). The gray dots represent the mean values for each participant and the red dots represent the mean values per relatedness condition (related, unrelated, semantically unrelated and control). Boxplots are shown in the violins. *p < .05. **p < .01.

Eye movement measures

Eye movement measures showed considerable effects of relatedness (see Figure 2C and Figure 2D). The ANOVA revealed a statistically significant effect of relatedness on number of fixations, $F(3,30) = 5.460$, $p = .004$, $\eta^2 = .201$, as well as on dwell time, $F(3,30) = 8.637$, $p < .001$, $\eta^2 = .250$. Post-hoc comparisons using the Tukey HSD test revealed that the number of fixations was significantly higher in the related condition (*M* = 0.490, *SD* = 0.800) compared to the control condition (*M* = 0.318, *SD* = 0.658), $t(40) = 2.809$, $p = .037$. This indicates that participants looked more often at the anchor than at the control distractor. Furthermore, post-hoc comparisons also revealed that dwell times were significantly higher in the related condition (*M* = 108.276 ms, *SD* = 201.109 ms) compared to the control condition (*M* = 59.341 ms, *SD* = 140.080 ms), $t(40) = 3.465$, $p = .007$. This indicates that participants looked longer at the anchor than at the control distractor. These results are in line with the hypothesis that related anchor objects are more distracting than random local objects when searching for a target object. Furthermore, as we hypothesized, we did not find significant differences in eye movements between the semantically unrelated condition and the control condition (see Table B1 and Table B2). This suggests that semantically unrelated anchor objects and random local objects are equally semantically related to the target object.

We could not provide significant evidence in favor of our other hypotheses as we found no significant differences in eye movements when we compared the other conditions with each other. However, as we expected, related anchor distractors and control distractors showed descriptively the highest difference in eye movements compared to the other comparisons. Moreover, when we look at the mean values in Figures 2C and 2D, we can descriptively see a trend in eye movements: The higher the relation between target and distractor object, the more distracting the distractor object was. These results indicate that hierarchical object-to-object relations seem to be reflected in our behavior.

Discussion

In the present study, we conducted an eye tracking experiment to investigate whether the presence of differently related anchor object affects the guidance of attention when searching for a target object on a search display. Particularly, we looked at differences in behavioral and eye movement measures due to differences in the relatedness between distractor objects (related, unrelated, semantically unrelated and control) and target objects. Although we did not find any significant effects on RTs, we found significant effects on number of fixations and dwell times. The significant differences between related anchor and control distractors in number of fixations and dwell times support our hypothesis that anchor objects which are highly semantically related to the target object are more distracting than random local objects when searching for a target object. Yet, we did not provide any significant evidence in favor of our other hypotheses assuming a trend in behavioral and eye movement measures.

With our study we aimed to combine two concepts with each other. On the one hand, we probed the hierarchical structure of a scene that Võ et al. (2019) proposed. On the other hand, we did this by cuing the target object with its function to make search natural and because phrases are formed based on objects' functions (Võ, 2021). In contrast to Castelhano and Witherspoon (2016) who studied effects of objects' functions on attention and visual search with invented objects, we used real-world objects. Besides, we did not prime the action by showing a video clip in which hands performed an action like Helbig et al. (2010). Instead, we used descriptions of the target object's function. To our knowledge, this has not been done in any study that has examined the role of anchor objects in visual search. Moreover, our study extends the work of Turini and Võ (2022) who studied the hierarchical structure of a scene by investigating the similarity of objects which belong to the same or different scene (or phrase or object type). In our study, we referred to the similarity as relatedness and investigated whether the relatedness had a distracting effect on action primed visual search.

We found that related anchor and control condition descriptively showed the highest difference in eye movements. Furthermore, we descriptively identified a trend suggesting that the higher the semantic relatedness between two objects, the stronger the distracting effect on eye movements. This trend is in accordance with Võ et al.’s (2019) proposed scene hierarchy as well as Turini and Võ’s (2022) findings about the similarity of objects, since it suggests, that the actions we primed by describing the target object’s function, activates a similarly structured network of representations. This implies that being confronted with an object’s function does not only activate a representation of the object itself, but also of the objects in its environment. Thus, this is in line with the concept of scene grammar (Biederman et al., 1982; Võ, 2021; Võ et al., 2019). Although we hypothesized that there would be no difference between semantically unrelated and control distractors, since both objects were semantically unrelated to the target object, we descriptively found a slight difference in eye movements. This could mean that anchor objects are fundamentally different from local objects. Thus, they could belong to a completely different category of objects. Usually anchor objects are larger and not easy to move, unlike local object, which are smaller and easy to move (Draschkow & Võ, 2017; Turini & Võ, 2022). This would implicate that our significant results were not due to our scene grammar and the hierarchical representation of Võ et al.'s (2019) proposed scene structure. Yet, in the inferential statistics, that difference was not significant.

In general, we assume that the trends we found descriptively are not significant, because our study with only 11 participants was underpowered. In the future, we would need more observations to increase power. Then, we might be able to see more effects on visual search behavior and significant results on within and between phrase comparisons. Besides, follow-up analysis could examine whether the salience (e.g., color) of the distractor objects has a distracting effect and therefore influenced the measures. Even though all pictures were reduced to the same size to keep low-level features constant, a *familiar-size Stroop effect* (Konkle & Oliva, 2012a) could have led to higher number of fixations and dwell times for anchor distractors.

A question that remains open is whether a virtual reality paradigm could provide more meaningful results than a search task on a circular display. Studies compared 2D searches with 3D searches and found that our memory improves when the whole body is involved in the search (Võ, 2021). One reason for this could be that in the real-world, we do not only use our gaze to interact with our surrounding. We physically interact with it (Draschkow & Võ, 2017). However, Võ (2021) pointed out that future research should try to find the right balance between tasks that resemble the real-world and controlled experiments. Another thing to consider is that, in the present study, we only used indoor scenes. That is why we cannot say whether the proposed hierarchical structure also applies to outdoor scenes. Future research could fill that gap.

To conclude, in the present study we combined knowledge of the hierarchical structure of a scene with the knowledge of object affordances. We found that hierarchical object-to-object relations that Võ et al. (2019) proposed seem to be reflected in action primed visual search. Therefore, we believe that the structure of a scene with its phrases plays an important role in cognitive processing of the world. With our study, we contributed to the understanding of visual cognition. However, further research with more participants is needed to gain more insights into the computation we need to exploit real-world regularities

References

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

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*(2), 143–177. https://doi.org/10.1016/0010-0285(82)90007-X

Boettcher, S. E. P., Draschkow, D., Dienhart, E., & Võ, M. L.-H. (2018). Anchoring visual search in scenes: Assessing the role of anchor objects on eye movements during visual search. *Journal of Vision*, *18*(13), 11. https://doi.org/10.1167/18.13.11

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Castelhano, M. S., & Witherspoon, R. L. (2016). How you use it matters: Object function guides attention during visual search in scenes. *Psychological Science*, *27*(5), 606–621. https://doi.org/10.1177/0956797616629130

Draschkow, D., & Võ, M. L.-H. (2017). Scene grammar shapes the way we interact with objects, strengthens memories, and speeds search. *Scientific Reports*, *7*(1), 16471. https://doi.org/10.1038/s41598-017-16739-x

Gibson, J. J. (1979). *The ecological approach to visual perception*. Psychology Press.

Helbig, H. B., Steinwender, J., Graf, M., & Kiefer, M. (2010). Action observation can prime visual object recognition. *Experimental Brain Research*, *200*(3–4), 251–258. https://doi.org/10.1007/s00221-009-1953-8

Henderson, J. M., Hayes, T. R., Rehrig, G., & Ferreira, F. (2018). Meaning Guides Attention during real-world scene description. *Scientific Reports*, *8*(1), 13504. https://doi.org/10.1038/s41598-018-31894-5

Konkle, T., & Oliva, A. (2012a). A familiar-size stroop effect: Real-world size is an automatic property of object representation. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 561–569. https://doi.org/10.1037/a0028294

Konkle, T., & Oliva, A. (2012b). A real-world size organization of object esponses in occipitotemporal cortex. *Neuron*, *74*(6), 1114–1124. https://doi.org/10.1016/j.neuron.2012.04.036

Lawrence, M. A. (2016). *Ez: Easy analysis and visualization of factorial experiments* (Version 4.4-0) [Computer software]. https://CRAN.R-project.org/package=ez.

Lenth, R. (2022). *Emmeans: Estimated marginal means, aka least-squares means* (Version 1.8.3). https://CRAN.R-project.org/package=emmeans

Mohr, J., Seyfarth, J., Lueschow, A., Weber, J. E., Wichmann, F. A., & Obermayer, K. (2016b). BOiS—Berlin object in scene database: Controlled photographic images for visual search experiments with quantified contextual priors. *Frontiers in Psychology*, *7*, 1–6. https://doi.org/10.3389/fpsyg.2016.00749

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, *11*(12), 520–527. https://doi.org/10.1016/j.tics.2007.09.009

Peirce, J. M., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Posit team. (2022). *RStudio: Integrated development environment for R* (Version 2022.12.0.353) [Computer software]. http://www.posit.co/

R Core Team. (2022). *R: A language and environment for statistical computing* (Version 4.2.2) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Turini, J., & Võ, M. L.-H. (2022). Hierarchical organization of objects in scenes is reflected in mental representations of objects. *Scientific Reports*, *12*(1), 20068. https://doi.org/10.1038/s41598-022-24505-x

Turk-Browne, N. B. (2012). Statistical learning and its consequences. In M. D. Dodd & J. H. Flowers (Eds.), *The Influence of Attention, Learning, and Motivation on Visual Search* (pp. 117–146). Springer New York. https://doi.org/10.1007/978-1-4614-4794-8_6

Võ, M. L.-H. (2021). The meaning and structure of scenes. *Vision Research*, *181*, 10–20. https://doi.org/10.1016/j.visres.2020.11.003

Võ, M. L.-H., Boettcher, S. E., & Draschkow, D. (2019). Reading scenes: How scene grammar guides attention and aids perception in real-world environments. *Current Opinion in Psychology*, *29*, 205–210. https://doi.org/10.1016/j.copsyc.2019.03.009

Võ, M. L.-H., & Wolfe, J. M. (2013). The interplay of episodic and semantic memory in guiding repeated search in scenes. *Cognition*, *126*(2), 198–212. https://doi.org/10.1016/j.cognition.2012.09.017

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis* (Version 3.4.0) [Computer software]. Springer-Verlag New York. https://ggplot2.tidyverse.org

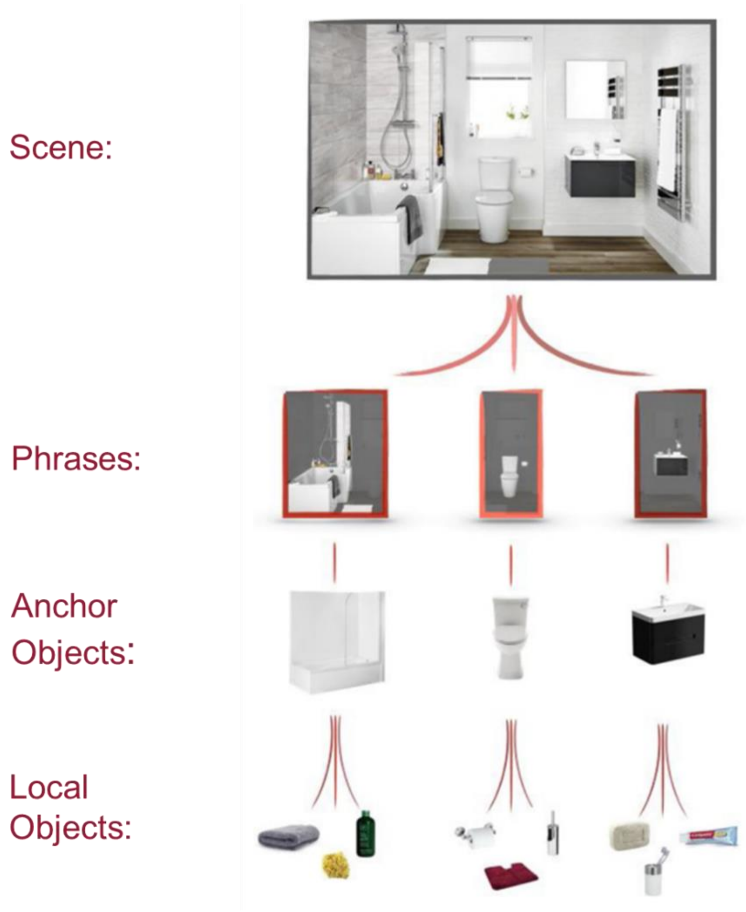
Wolfe, J. M. (1994). Guided search 2.0. A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. https://doi.org/10.3758/BF03200774

Wolfe, J. M., Võ, M. L.-H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, *15*(2), 77–84. https://doi.org/10.1016/j.tics.2010.12.001

Appendix A

Figure A1

Proposed Scene Hierarchy of a Bathroom Scene



Note. This figure describes the current opinion in psychology (adapted from Võ et al., 2019, p. 207). The bathroom scene consists of three phrases with one anchor object and several local objects each.

Figure A2

Exemplary Objects with different Object-to-Object Relations

Note. This figure shows one exemplary distractor object for each condition (related, unrelated, semantically unrelated and control) that are differently related to the target objects.

Appendix B

Table B1

Post-hoc Comparisons for Number of Fixations

Contrast	Estimate	SE	df	t.ratio	p-value
Related - Unrelated	0.052	0.059	40	0.882	.814
Related - Semunrelated	0.138	0.059	40	2.326	.109
Related - Control	0.167	0.059	40	2.809	.037*
Unrelated - Semunrelated	0.086	0.059	40	1.444	.480
Unrelated - Control	0.114	0.059	40	1.927	.233
Semunrelated - Control	0.029	0.059	40	0.483	.962

Note. N = 11. Here you can see the results of the post-hoc comparisons with the Tukey HSD test for number of fixations. *p < .05.

Table B2

Post-hoc Comparisons for Dwell Time (ms)

Contrast	Estimate	SE	df	t.ratio	p-value
Related - Unrelated	24.7	13.8	40	1.789	.294
Related - Semunrelated	37	13.8	40	2.679	.051
Related - Control	47.9	13.8	40	3.465	.007*
Unrelated - Semunrelated	12.3	13.8	40	0.890	.810
Unrelated - Control	23.2	13.8	40	1.676	.349
Semunrelated - Control	10.9	13.8	40	0.786	.860

Note. N = 11. Here you can see the results of the post-hoc comparisons with the Tukey HSD test for dwell time. *p < .05.

Statutory Declaration

I herewith declare that I have composed the present student paper myself and without use of any other than the cited sources and aids. Sentences or parts of sentences quoted literally are marked as such; other references with regard to the statement and scope are indicated by full details of the publications concerned. This student paper in the same or similar form has not been submitted to any examination body and has not been published. It was not yet, even in part, used in another examination or as a course performance.

Place, Date: Frankfurt, ____________ Signature: ____________