Time course and shared neurocognitive mechanisms of mental imagery and visual perception

Martin Maier [1,2], Romy Frömer [1,3], Johannes Rost [1], Werner Sommer [1,2], and Rasha Abdel Rahman [1,2]

[1] Department of Psychology, Humboldt-Universität zu Berlin
[2] Berlin School of Mind and Brain, Humboldt-Universität zu Berlin
[3] Department of Cognitive, Linguistic, and Psychological Sciences, Brown University

Running head:	Mental imagery: time course, cognitive mechanisms
Address for correspondence:	Rasha Abdel Rahman, Humboldt-Universität zu Berlin, Rudower Chaussee 18, 12489 Berlin, Germany
Phone:	+49-(0)30-2093-9413
Email:	rasha.abdel.rahman@hu-berlin.de
Word count:	Main text: 3912, Methods: 1767

Acknowledgements

This work was funded by the German Research Foundation grant AB 277-6 to Rasha Abdel Rahman. Martin Maier was supported by the state of Berlin with an Elsa Neumann scholarship and by the Berlin School of Mind and Brain. We thank Rainer Kniesche for technical assistance.

Abstract

When we imagine an object and when we actually see that object, similar brain regions become active. Yet, the time course of neurocognitive mechanisms that support imagery is still largely unknown. The current view holds that imagery does not share early perceptual mechanisms, but starts with high-level visual representations. However, evidence of early shared mechanisms is difficult to obtain because imagery and perception tasks typically differ in visual input. We therefore tracked electrophysiological brain responses while fully controlling visual input, (1) comparing imagery and perception of objects with varying amounts of associated knowledge, and (2) comparing the time courses of successful and incomplete imagery. Imagery and perception were similarly influenced by knowledge already at early stages, revealing shared mechanisms during low-level visual processing. It follows that imagery is not merely perception in reverse; instead, both are active and constructive processes, based on shared mechanisms starting at surprisingly early stages.

Keywords: mental imagery, early visual processing, event-related potentials, semantic knowledge, P1 component

Time course and shared neurocognitive mechanisms of mental imagery and visual perception

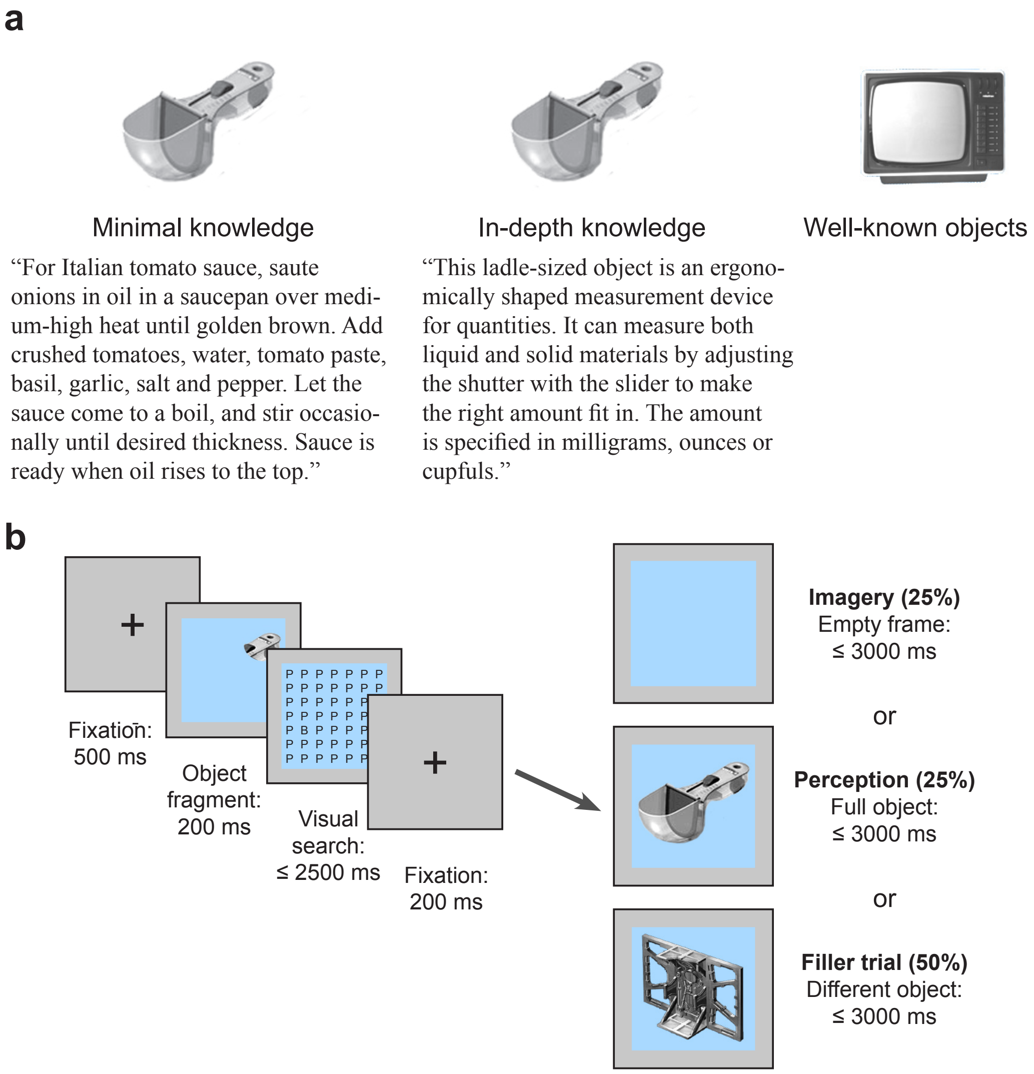
A growing body of research suggests that seeing something with the mind's eye—mental imagery—may not be all that different from seeing something with one's physical eyes. Indeed, imagery and perception recruit overlapping neural circuits, including primary visual areas [1-8], and the vividness of imagination correlates with the similarity of brain activities accompanying imagery and perception [9].

Predictive processing accounts posit that perception arises from hierarchical Bayesian predictions—essentially imaginations—that are constrained by bottom-up sensory input [10-14]. This theoretical framework is neurally plausible [15-20] and supported by evidence that even early stages of perception are subject to top-down influences [15,21-31]. This suggests that initial aspects of imagery could be fast enough to generate early top-down effects.

This suggestion contrasts with alternative accounts assuming that perception first runs through a strictly hierarchical succession of increasingly complex visual representations, with early stages mainly driven by bottom-up sensory processes. At later stages, recurrent feedback from higher-level brain areas is assumed to enable stabilization of visual representations and, eventually, conscious access [32,33]. Based on this account of perception, recent work has mapped out how visual imagery could follow a reverse hierarchy of activation compared to perception [6,7,34-36]. Under these assumptions, imagery would not rely on early perceptual mechanisms like feature processing but start relatively late, with entire visual representations that bring several levels of the visual hierarchy into concert [6,7,35]. In support of this idea, Dijkstra, et al. [35] found neural activation patterns during imagery to correspond to those found during high-level perception, but not early, low-level stages of perception. Earlier studies reported imagery-related variations in the N1 component of the event-related potential (ERP) [37-39] that is associated with configural visual processing [40-44]. Here, we refer to configural visual processing as the encoding of constituent features into meaningful configurations (e.g., whole objects) [45]. Yet, the designs commonly used to compare perception and imagery are not optimal for providing evidence of shared mechanisms during early visual processing because imagery and perception conditions often involve substantially different visual stimulation [35,36,38]. This may mask early common neural mechanisms, in particular, given that brain activity in early visual processing is more strongly influenced by low-level stimulus properties [32,33]. Here we propose a way to overcome this obstacle by varying the content of imagery while controlling for visual properties. This allows us to compare the time course and functional mechanisms of imagery and perception from initial to final stages and, specifically, to test for parallels at earlier stages than previously reported. If so, we would have to revise our current understanding of the mechanisms supporting mental imagery and how they unfold over time.

Our approach borrows from designs used in perception research to investigate changes in early visual processing independent of the specific visual input [24,27,46]. This is achieved by manipulating the semantic knowledge associated with a given object. Knowledge stored in semantic memory, for example, about the functions of objects [24,27,31,46], and categories defined by the language we speak [25,26,28,47-49], have all been shown to influence early visual processes.

We combined this approach with recording and analyzing ERPs to test with high temporal precision whether early top-down effects, repeatedly observed in perception, are mirrored in imagery. Based on previous findings [24,27,46], we expected semantic knowledge to decrease the P1 component in the ERP, a marker of sensory processing sensitive to low-level visual features such as luminance and contrast [16,50-53], as well as the later N400 component, reflecting high-level semantic processing [24,27,54]. Crucially, we predicted that knowledge would influence both components similarly in perception as well as imagery.



"For Italian tomato sauce, saute onions in oil in a saucepan over medium-high heat until golden brown. Add crushed tomatoes, water, tomato paste, basil, garlic, salt and pepper. Let the sauce come to a boil, and stir occasionally until desired thickness. Sauce is ready when oil rises to the top."

"This ladle-sized object is an ergonomically shaped measurement device for quantities. It can measure both liquid and solid materials by adjusting the shutter with the slider to make the right amount fit in. The amount is specified in milligrams, ounces or cupfuls."

Figure 1. Study design. (a) Knowledge conditions with examples of object-unrelated information (minimal knowledge condition) and object-related information (in-depth knowledge condition). (b) Trial types and structure of the main task. All trial types came in all knowledge conditions (minimal, in-depth and well-known), with equal probability and in randomized order.

We further compared successful and incomplete imagery, akin to previous studies leveraging vividness ratings [7,9], to determine the processing stages that drive successful imagery without confounding influences from visual input. We assumed that successful and incomplete imagery would show similar activation patterns during low-level visual processing (P1 component), but would differ in high-level, configural visual processing (N1 component).

Finally, to gain further understanding of the mechanisms driving mental imagery, we tested how the neural dynamics that dissociate successful and incomplete imagery relate to perception.

Results

To investigate whether perception and imagery rely on shared early perceptual mechanisms and examine their time course, we recorded EEG from 32 participants while they viewed or imagined objects with varying amounts of associated knowledge. Target objects were cued with object fragments and, following an intervening visual search task to reset visual activity, participants either made a familiarity judgment on a presented object or imagined the cued object on an empty frame (see Figure 1).

Behavioral results

In the imagery task, participants were asked to form intact and detailed mental images of the cued objects. They indicated successful and incomplete imagery via button press. Overall, participants indicated successful imagery in 84.5 % of the trials. Imagery success rates were higher in the well-known compared to the in-depth knowledge condition (89.2 % vs. 83.0 %; nested binomial GLMM: $b = 0.53$, $z = 5.58$, $p < .001$), but there was no difference between the in-depth and the minimal knowledge condition (83.0 % vs. 81.1 %; $b = 0.08$, $z = 0.88$, $p = .380$; see Figure 2). Knowledge affected reaction times (RTs), which gradually decreased with the depth of knowledge, indicating faster imagery for objects learned with in-depth compared to minimal knowledge (1730.4 vs. 1770.2 ms; $b = -0.02$, $t = -2.04$, $p = .042$), and for well-known objects compared to objects with in-depth knowledge (1673.7 vs. 1730.4 ms; $b = -0.05$, $t = -3.06$, $p = .003$).

In the perception task, participants classified object pictures as newly learned vs. well-known. Classification accuracy was lower in the well-known compared to the in-depth knowledge condition ($b = -0.88$, $z = -4.89$, $p < .001$) and also in the in-depth compared to the minimal knowledge condition ($b = -0.55$, $z = -2.23$, $p = .026$). RTs in the perception task (Figure 2) did not differ across knowledge conditions (nested LMM, in-depth - minimal: $b < 0.01$, $t = .61$, $p = .544$; well-known - in-depth: $b = -0.02$, $t = -.90$, $p = .375$). Lower accuracy in classifying well-known objects can be explained by context effects: Participants were to classify well-known objects as “old”, but these objects had been rare during the learning session, thus in the context of the test session they were “new”. In contrast, participants were to classify newly learned objects as “new”, but in the context of the experiment these objects had been seen many times, and objects associated with richer semantic knowledge may have seemed subjectively more familiar, and thus “old”. While the incongruence between long-term semantic knowledge and contextual familiarity may have muddied the waters, the observation of facilitated imagery demonstrates that our semantic knowledge manipulation was effective.

Effects of semantic knowledge on ERPs

To test the hypothesis that imagery and perception share knowledge-related modulations of early visual activity, we analyzed the effects of semantic knowledge on the P1 component, an index of early perceptual processing. We further tested for later effects of knowledge in the N400, an indicator of semantic processing. In line with our hypothesis, across both imagery as well as perception, P1 amplitudes decreased with semantic knowledge, yielding significant reductions from minimal to in-depth, and from in-depth knowledge to well-known objects (Figure 2, Table 1). The full LMMs, including semantic knowledge, task and their interactions revealed no significant interactions of knowledge and task, suggesting similar effects of knowledge in both conditions. Exclusion of these interactions further did not significantly decrease model fit, $\Delta X^2(4) < 5.19$, $p > .268$, and fit indices favored the reduced models (ΔAIC$_{P1}$: -7, ΔBIC$_{P1}$: -37; ΔAIC$_{N400}$: -3, ΔBIC$_{N400}$: -32). In the N400, well-known objects produced significantly more negative amplitudes than newly learned objects, whereas the minimal and in-depth knowledge conditions did not differ.

Given differences in visual stimulation, trivial differences in ERP amplitudes between the tasks are expected. Indeed, across both ERP components, we found more positive amplitudes for perception, while there was no difference between imagery and incomplete imagery.

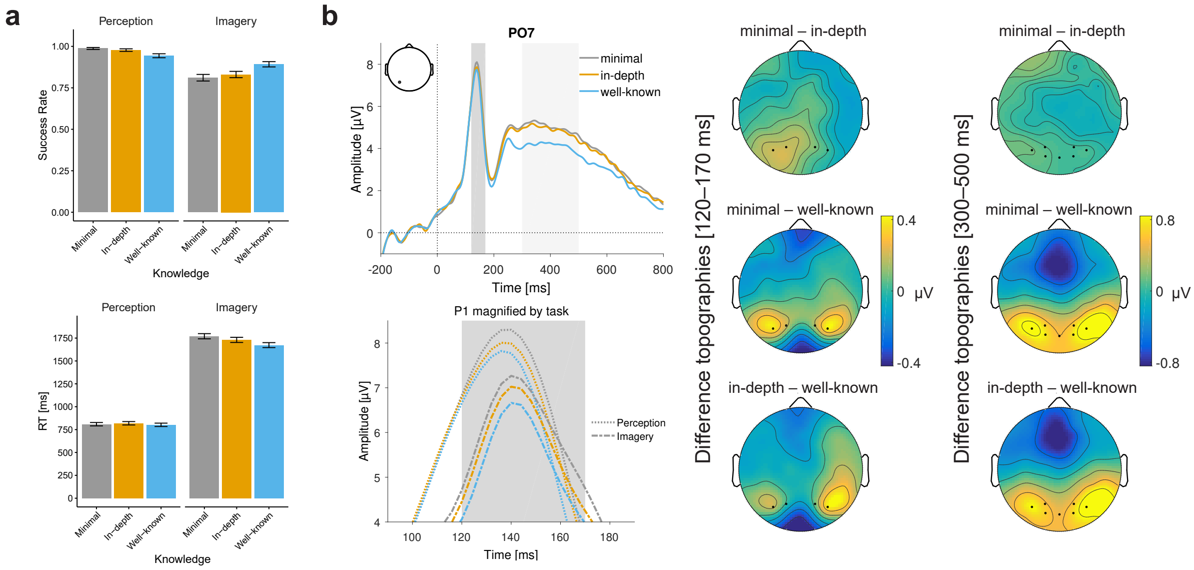


Figure 2. Semantic knowledge effects. (a) Behavioral results: Accuracy in the perception task and imagery success rate (top) and mean RTs (bottom) as a function of object knowledge. Error bars represent 95% confidence intervals. (b) Effects of object knowledge on the P1 and N400 components. Left, top panel: Grand average ERPs at electrode PO7, aggregated over perception and imagery. Bottom panel: Zooming in on the P1 peak illustrates comparable knowledge effects in imagery and perception. Right panel: Difference topographies comparing the knowledge conditions in the P1 and N400 time windows (120-170 ms; 300-500 ms, respectively). Region of interest (ROI) electrodes are marked as dots.

Table 1

Knowledge effects on the P1 and N400 components during perception and mental imagery.

	P1 amplitude				**N400 amplitude**			
Predictors	*Estimates*	*SE*	*t-value*	*p-value*	*Estimates*	*SE*	*t-value*	*p-value*
(Intercept)	5.09***	0.61	8.38	<0.001	3.65***	0.42	8.64	<0.001
Visual (Ima–Per)	-0.93*	0.39	-2.39	0.016	-0.71*	0.35	-2.04	0.042
Visual (Nima–Ima)	0.28	0.16	1.72	0.086	0.30	0.21	1.46	0.146
Knowledge (Deep–Min)	-0.26*	0.10	-2.55	0.011	-0.17	0.11	-1.50	0.133
Knowledge (Well–Deep)	-0.30*	0.14	-2.12	0.033	-0.45**	0.17	-2.61	0.009
Random Effects	*SD*				*SD*			
Participants	3.41				2.34			
Visual (Ima–Per)	2.14				1.90			
Visual (Nima–Ima)	0.14				0.57			
Object Identity	0.35				0.47			
Residual	4.27				4.70			
Deviance	61305.513				63272.899			
log-Likelihood	-30652.756				-31636.449			

Note: Visual (Ima–Per) = Perception – Imagery, Visual (Nima–Ima) = Incomplete Imagery – Imagery, Knowledge (Deep–Min) = in-depth – minimal, Knowledge (Well–Deep) = well-known – in-depth
* $p<0.05$ ** $p<0.01$ *** $p<0.001$

Knowledge effects on the P1 in perception have been repeatedly observed in the absence of cueing [24,27,46], and any visual priming in the present study could only occur partially as we only showed object fragments followed by an intervening visual search task to reset visual activity. Nevertheless, a potential remaining concern in the current design is that knowledge effects on the P1 may reflect spillovers from the cues. If this were true, we should observe knowledge effects also on filler trials, where non-cued objects were shown. In a control analysis we found no evidence that the knowledge condition of the object cue influenced the P1 in filler trials. There was no difference between the well-known and the in-depth knowledge condition (LMM$_{\text{Fillers}}$: $b = -0.070$, $t = -0.674$, $p = .500$) or between the in-depth and the minimal knowledge condition (LMM$_{\text{Fillers}}$: $b = 0.002$, $t = -0.021$, $p = .984$).

Thus, knowledge effects in the P1 appear to be specific to imagining or seeing the corresponding objects.

To summarize, in line with our hypothesis we found semantic knowledge effects in early visual processes across both imagery and perception: P1 amplitudes were reduced with increasing depth of object-related knowledge. This effect replicates previous findings from visual perception[24,27,46] and extends them to imagery. Previously reported differences between minimal and in-depth conditions in the N400, reflecting high-level semantic processes, were not replicated [24,27].

Comparisons between successful and incomplete imagery

To better understand the mechanisms that differentiate between successful and unsuccessful imagery, we compared trials in which participants had indicated the former vs the latter. The hypothesis was that incomplete compared to successful imagery may arise from failed configural processing and should thus be associated with differences in the N1 component. Since imagery may be supported by increased fronto-posterior coupling [34,55], differences in frontal activity were also expected. Even though EEG scalp distributions do not translate easily to generators of activity in the brain, we hypothesized that posterior N1 effects may therefore coincide with mirrored effects at frontal sites [56]. To test for global differences between successful and incomplete imagery we compared mean amplitudes with the cluster-based permutation test approach (CBPT), which revealed a significant difference. Underlying this difference were two clusters across electrodes and time: a posterior cluster between 228 and 392 ms, and a frontoparietal cluster between 304 and 492 ms that was slightly lateralized to the right hemisphere (Figure 3). As expected, the beginning and topography of the posterior cluster suggested a modulation of the N1 component (Figure 3).

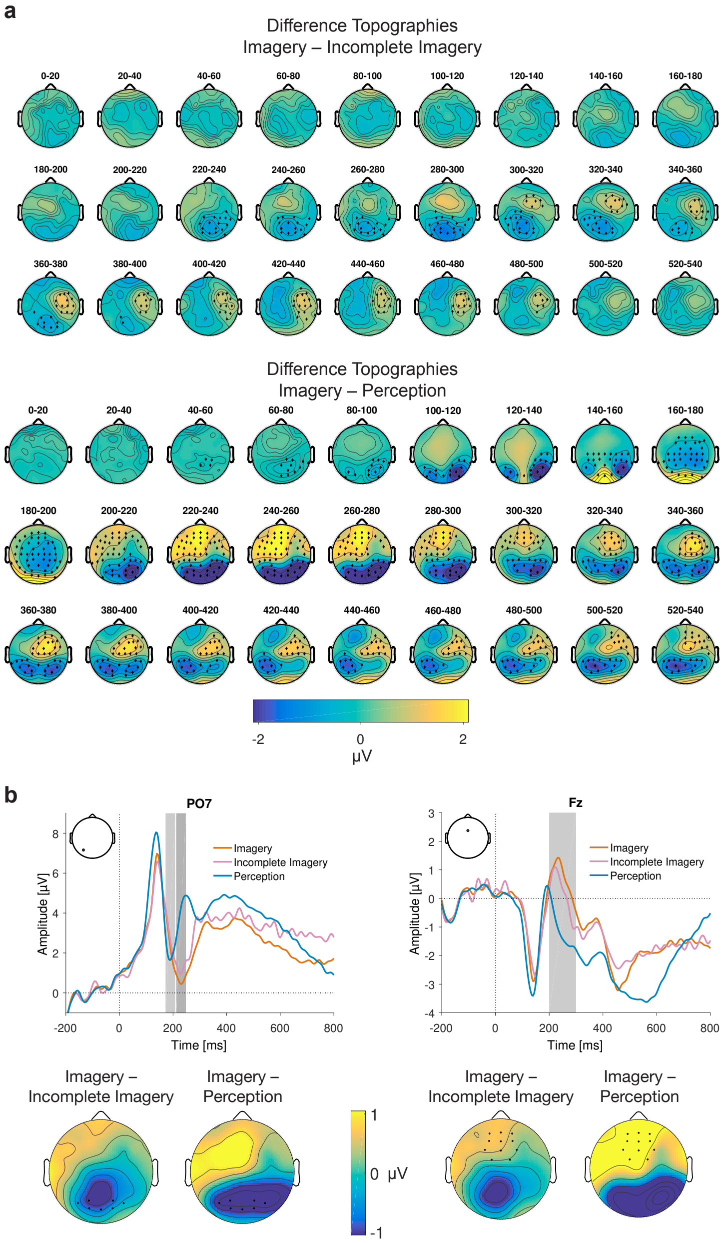


Figure 3. ERP-comparisons of successful imagery, incomplete imagery, and perception. (a) Grand average difference-topographies. Highlighted electrodes are part of spatio-temporal clusters most compatible with the significant differences between successful and incomplete imagery (top), and between imagery and perception (bottom). (b) Comparisons of successful imagery, incomplete imagery, and perception in the N1 time window. Time windows entering the analysis of the posterior N1 amplitudes (left) and the simultaneous frontal positivity (right) are highlighted with grey shading. Topographies illustrate differences in the highlighted time windows with ROI electrodes marked by dots.

Follow-up LMM analyses based on single trial amplitudes in an independently determined posterior ROI (see Method) confirmed a significant difference in the N1 component. Successful imagery was characterized by a larger N1 compared to incomplete imagery (Table 2). Around the same time, successful and incomplete imagery also differed at frontal sites, with a larger positivity in the frontal ROI in successful imagery trials (Figure 3, Table 2). Thus, the comparison between successful and incomplete imagery aligns with our hypothesis that successful imagery is supported by mechanisms of configural processing indexed by the posterior N1 and potentially supported by frontal top-down regulation.

To test whether the same neural dynamics dissociate between imagery and perception, we compared these conditions using the same two-step approach. CBPT revealed significant differences between perception and imagery. Starting with a relative negativity for imagery at parieto-occipital sites around 80 ms post stimulus, all remaining time windows yielded significant clusters (cf. Figure 3). As outlined above, early differences between imagery and perception are trivial due to differences in visual stimulation. Further, differences between imagery and perception could be driven by latency shifts, amplitude differences, or both. We therefore analyzed peak latencies of key ERP components—P1 and N1—in the different visual conditions (perception, successful and incomplete imagery as one factor). P1 and N1 peak latencies were detected in the average ERP at PO7 for each participant and condition. Indeed, latency of the posterior N1 component was significantly delayed by an estimated 27 ms in imagery compared to perception ($\text{LMM}_{\text{Imagery-Perception}}$: $b = 27.75$; $t = 2.88$; $p = .005$), while there were no reliable latency shifts in the P1 component ($\text{LMM}_{\text{Imagery-Perception}}$: $b = 5.87$; $t = 1.62$; $p = .110$).

Table 2

Comparisons between imagery, incomplete imagery, and perception in the N1 time window.

	N1 amplitude				**Frontal P1 amplitude**			
Predictors	*Estimates*	*SE*	*t-value*	*p-value*	*Estimates*	*SE*	*t-value*	*p-value*
(Intercept)	1.66 ***	0.36	4.57	<0.001	0.20	0.22	0.89	0.383
Visual (Ima–Per)	-0.96 *	0.42	-2.27	0.030	0.62 **	0.21	2.92	0.006
Visual (Nima–Ima)	0.54 ***	0.15	3.63	<0.001	-0.24 *	0.12	-2.04	0.041
Centered P1	4.20 ***	0.05	79.76	<0.001				
Centered N1f					0.71 ***	0.01	81.06	<0.001
Random Effects	*SD*				*SD*			
Participants	2.03				1.22			
Visual (Ima–Per)	2.36				1.12			
Object Identity	0.27				0.12			
Visual (Ima–Per)					0.31			
Residual	3.94				3.12			
Deviance	59540.162				54515.483			
log-Likelihood	-29770.081				-27257.741			

Note: Visual Ima–Per = Perception – Imagery, Visual Nima–Ima = Incomplete Imagery – Imagery, P1 = preceding posterior P1 component, N1f = preceding frontal N1 component
* $p<0.05$ ** $p<0.01$ *** $p<0.001$

The LMM analysis of N1 amplitudes was adjusted for these latency shifts (time windows are highlighted in Figure 3). To account for the differences in visual stimulation between imagery and perception, we further included centered trial-by-trial P1 amplitudes as a covariate. This can be seen as a kind of baseline correction [57] because the P1 should capture a large portion of the variance related to differences in visual input and correct for amplitude differences resulting from evoked amplitude variance. When testing for an interaction between P1 amplitude and visual condition it was not significant. Exclusion of the interaction did not significantly decrease model fit, $\Delta X^2(2) = 3.28$, $p = .194$, and fit indices favored the reduced models (ΔAIC: -1, ΔBIC: -15). The N1 was significantly larger for successful imagery compared to perception (Table 2). The N1 further increased (became more negative) with more positive P1 amplitudes. Thus, the difference between perception and imagery found in the overall CBPT analysis appears to be driven by latency *and* amplitude differences.

Like for the comparison between successful and incomplete imagery, there was a modulation at frontal sites, where we found a larger positivity for imagery compared to perception coinciding with the posterior N1 component (Figure 3, Table 2). The frontal P1 further increased with more positive amplitudes of the preceding frontal negativity, which we controlled in order to account for earlier visually evoked differences.

To summarize, we found a larger posterior N1 for successful compared to incomplete imagery and for imagery compared to perception. These effects were accompanied by modulations of a frontal positivity in the approximate time range of the N1, which was significantly enhanced for successful compared to incomplete imagery, as well as for imagery compared to perception. Taken together these findings indicate increased demands on configural processing in imagery compared to perception, potentially supported by increased recruitment of frontal top-down processing, and that imagery fails if these increased demands are not met.

Discussion

It is now widely accepted that visual perception and mental imagery rely on shared brain circuits, including regions in early visual cortex, as well as frontal and parietal regions [2,7]. Yet, the time course of imagery and the timing of the involvement of early visual cortex are still open questions. In line with predictive processing accounts one hypothesis holds that perception engages top-down predictions even during low-level processing [25,26,31,47], and that imagery might share this mechanism.

A different hypothesis based on a more strictly hierarchical account of perception is that imagery works like perception in reverse, assuming that it activates the entire visual representation from the start, and does not rely on early perceptual representations [6,7,34,35].

This account is supported by work showing similarities of brain activity between imagery and high-level perception [35], and by imagery-related effects at the level of configural processing, as reflected in the N1 component of the ERP [37-39,58]. Thus, late involvement of early visual areas is mainly supported by a lack of evidence for early involvement. Such evidence is difficult to obtain, however, when the visual input between imagery and perception differs [35,36].

To overcome this obstacle, we varied the amount of knowledge associated with objects that participants saw and imagined—that is, we manipulated top-down predictions while keeping bottom-up input constant. This allowed us to detect changes in early visual activity independent of the visual stimulation. Using this approach, we show that like in perception, semantic knowledge modulates early visual activity also during imagery, revealing similar mechanisms at a much earlier stage than previously assumed. We further show that successful imagery is characterized by increased activity during high-level, configural visual processing compared to both, incomplete imagery as well as perception. This suggests that demands on configural processing are higher in the absence of supporting bottom-up input, and that rather than initiating imagery, stable visual representations need to be constructed, much like in perception.

Knowledge facilitates imagery and shapes early stages of imagery and perception

In imagery, like in perception, object-related knowledge and familiarity influenced visual processing at an early stage. Deeper knowledge was associated with decreases in the amplitude of the P1 component that reflects low-level visual processing in extrastriate visual areas [16,50,51]. If knowledge can influence imagery at this stage, it suggests that at least some imagery-related processes take place in early visual areas already at an early latency. Object knowledge appears to inform top-down predictions that are used in both, imagery and perception. The effect being located in the P1 component demonstrates an influence on the processing of low-level object features. We conclude that knowledge about an object’s function and its relevant parts facilitates low-level feature processing when we see or when we imagine an object. These findings demonstrate that imagery and perception rely on shared top-down mechanisms in the construction of low-level visual representations.

Notably, the influence of semantic knowledge on imagery was of direct behavioral relevance: imagery of well-known objects was more often successful and faster than imagery of less familiar objects. Additionally, imagery was faster when participants had acquired in-depth rather than only minimal knowledge about initially unfamiliar objects. Thus, the more we know about an object, the better we can imagine it.

While the P1 component in the imagery condition was evoked by a visual stimulus, the presentation of a light blue square, this physical stimulus was identical for the semantic knowledge conditions and can, hence, not have produced the observed knowledge effects. A potential objection is that modulations of early visual ERPs might not have been related to imagery, but to spillovers from the object fragment cue. However, this explanation is unlikely as 1) the same semantic knowledge effects on perception have been shown in the absence of cueing [24,27,46], 2) we only presented fragments of the objects to be imagined or perceived, 3) visual input was reset by an intervening visual search task, and 4) there were no cue-related knowledge effects for filler trials. We therefore conclude that object knowledge influences low-level visual processes during both, perception and imagery.

At variance with our predictions, we did not observe an influence of in-depth versus minimal semantic knowledge on the N400 component. In contrast to previous studies demonstrating these effects in perception [24,27], here, we cued the objects, which likely triggered object recognition and semantic processing. Whereas the intervening visual search task interfered with visual working memory, higher-level semantic network activation of the current object might have been sustained, given that it was potentially relevant for the upcoming task. Since the N400 is typically smaller for expected stimuli and reflects changes in semantic network activation [59], the cues in our paradigm may have muted the N400 effects.

What distinguishes successful imagery from incomplete imagery and from perception?

In line with our hypothesis that imagery relies strongly on configural visual processing, successful and incomplete imagery started to diverge in the posterior N1 component [40-44]. Successful imagery was associated with larger posterior N1 amplitudes accompanied by larger frontal positive amplitude modulations. The former finding is consistent with previous EEG and MEG studies that showed imagery-related modulations of the posterior N1 [37-39,58]. In terms of its functional relevance and typical latency, the N1 effect fits well with the finding that neural representations decoded from imagery using MEG match those observed in perception around 160 ms, that is, the N1 time window [35]. As incomplete imagery did not differ from successful imagery in the P1, it seems to share the early low-level activations but to lack (some of) the later configural processes and top-down feedback that stabilizes the image. The reduced frontal activity may thus reflect insufficient involvement of frontal areas, and their connectivity to occipitotemporal visual areas, which provide crucial top-down monitoring for imagery to be maintained [7,34,55]. Holding intact and detailed images before the mind's eye thus seems to be supported by configural visual processing and large-scale connectivity including frontal and occipital areas that stabilizes and maintains visual representations [6,34,55].

This interpretation is further supported by our findings comparing imagery and perception. We found that the posterior N1 was both delayed and increased in imagery compared to perception. Simultaneously, frontal activity was more pronounced in imagery than in perception. These results suggest that imagery relies more strongly on configural processing than perception, and engages more top-down control. When these additional demands are not met, imagery fails. To test whether success vs failure of imagery is all or none or reflects gradual degradation in configural processing, future studies could employ trial-by-trial vividness ratings to test whether these correspond to linear decreases in frontal and posterior activity [9].

Taken together, across perception and imagery, we found modulations of early visual processing by semantic knowledge. Compared to both perception and incomplete imagery, successful imagery was characterized by increased frontal and posterior activity in the N1 time range, presumably reflecting increased connectivity between higher level control and lower level visual areas to support configural processing.

Interestingly, this pattern bears similarities to what we know about conscious access. The P1 component is not typically associated with perceptual awareness [26,60], and also did not dissociate between successful and incomplete imagery in the present study. Conscious perception is thought to depend on “global ignition” or recurrent processing in a widespread network of brain areas [32,61]. It is therefore conceivable that differences between successful and incomplete imagery starting in the N1, as well as late, high-level visual representations decodable around 500 ms [35], reflect the beginning of *conscious* mental imagery, not the beginning of imagery-related processing per se. The earlier imagery-related processing stages revealed by knowledge effects on the P1 could be pre-conscious, just as in perception.

What we learn about perception

The fact that we find the same knowledge effects on the P1 in imagery and perception also teaches us something relevant about perception. Recently, the debate if there are any true top-down effects on perception has sparked new controversy [25,62-64]. Here we show semantic top-down influences on early visual processing in the absence of the relevant physical stimulus. This demonstrates that knowledge can have true top-down effects on early and automatic stages of perception. This is in line with the predictive processing account in which perception is seen as a process of active hierarchical Bayesian inference [10-14]. It construes perception more from the inside out than from the outside in: what we perceive is described as the brain’s best guess about the causes of afferent sensory input. The fact that imagery and perception appear to share early top-down predictions brings to mind the notion that perception might be a form of “controlled hallucination” [10]. Perception might actually have elements of controlled imagery—involving a form of non-voluntary and pre-conscious imagery that is triggered and constrained by sensory input [65].

Conclusion

Our results provide important insights into the time course of visual mental imagery by demonstrating that top-down influences modulate imagery already at an early stage of low-level visual feature processing. This challenges the idea that imagery and perception share neural substrates only for high-level visual processes. Instead, they engage common neurocognitive mechanisms already during early visual processing stages— consisting in top-down predictions, informed by knowledge stored in memory. Whether in seeing or imagining objects, our brains begin to construct what we "see" before the mind's eye from basic visual features and with the help of what we know.

Methods

Participants

Participants were 32 native German speakers (23 women; mean age 24 years; age range 20-35). All were right-handed with normal or corrected-to-normal visual acuity. Two participants were replaced due to excessive EEG artifacts. The study was approved by the Ethics Committee of the Humboldt-Universität zu Berlin. Participants gave written informed consent and received payment or course credits.

Apparatus and stimuli

Stimuli were presented on a 17'' monitor using Presentation (Neurobehavioral Systems ®, Berkeley, USA) with a viewing distance of approximately 90 cm. The stimulus set comprised 40 rare objects [24] unfamiliar to all participants (Figure 1) and 20 well-known objects. All stimuli were gray-scale pictures of either entire objects or object fragments (used as cues), covering about 20% of the object, all displayed on a blue background frame of 3.5 × 3.5 cm (2.22° × 2.22° visual angle; see Figure 1). Object fragments were typical parts of the corresponding objects, allowing recognition. Fragment positions (center, left, right, top, bottom part of the object) were counterbalanced across objects. During learning, object names consisting of pseudo-nouns uninformative regarding the object's function, were presented in both written and spoken form. In addition, for each unfamiliar object, an audio description was presented containing either a short explanation of the object's function, use and origin (mean duration 18.3 s), or a cooking recipe (out of 20 recipes; mean duration 18.6 s, see Figure 1).

Visual search displays consisted of a 7 by 7 matrix of uppercase letters with one single deviant letter (see Figure 1). One of three different letter combinations (F-E, P-B, and T-L) was shown on a light blue background measuring 5 × 3.5 cm (3,17° × 2.22°). The deviant letter could appear in any position of the matrix except for the center column.

Task and procedure

All participants completed two sessions on different days: a learning session, in which they acquired semantic knowledge about unfamiliar objects, and a test session that tested imagery and perception of the learned objects along with well-known objects.

Learning Phase. The learning session consisted of two parts. In Part 1, lasting about 45 minutes, participants were presented with 40 unfamiliar objects and their names (written and spoken). The first part ended with a short test (approximately 10 min), comprising verbal naming and familiarity decisions on both, well-known objects and newly learned objects.

In Part 2, lasting about 75 min, participants listened to recordings that provided object-related information about origin, function and use of half of the unfamiliar objects (in-depth knowledge condition), and unrelated cooking recipes for the other half (minimal knowledge condition). Object–knowledge combinations were counterbalanced across participants, such that each object was equally often part of both knowledge conditions. All stories were presented twice. Thus, all unfamiliar objects were presented equally often and for the same duration and only object-related knowledge was manipulated. This resulted in three conditions with increasingly elaborate knowledge: newly learned objects without functional information (20 objects, minimal knowledge condition), newly learned objects with detailed information (20 objects, in-depth knowledge condition), and well-known objects, with preexisting information, visual and hands-on experience (20 objects, well-known objects condition). Part 2 ended with the same naming and familiarity test as Part 1.

Test Phase. The test session, which included EEG recordings, took place two to three days after the learning session. Before the experiment, participants filled in a knowledge questionnaire, testing recall of the pictures and related information of newly learned and well-known objects. Then, they were familiarized with the object fragments, to make sure they could recognize the corresponding objects. Before the main task, participants performed a practice block with five well-known objects (not part of the test set), which was repeated up to two times if necessary.

In the main task, participants either imagined or saw pictures of objects. Investigating imagery with ERPs bears some timing-related difficulties: the content of imagery must be cued, but cue processing should not overlap in time with imagery, and the precise onset of imagery should be controlled. Furthermore, effects of object-knowledge on neural processing should be related to imagery, not processing of the cue. We designed a task to control the onset and content of imagery (Figure 1). First, an object fragment was presented as cue, followed by a demanding visual search trial meant to delay the onset of imagery by taxing visual working memory [66-68], and as a precaution against the transfer of semantic effects induced by the cue to the onset of mental imagery. Participants were instructed to indicate the position of a deviant letter in the left or right half of the display. Next, participants either saw an empty blue frame (imagery task, 180 trials, 25 %), a full picture of the cued object (perception task, 180 trials, 25 %), or a different object (filler trials, 360 trials, 50 %). Immediately after a response or if no response had been given within 3 s after stimulus onset, a blank screen of 1 s duration was presented.

In the imagery task, participants were instructed to form an intact and detailed mental image of the cued object as quickly and accurately as possible. Participants indicated successful or incomplete imagery via button press. In perception and filler trials, participants indicated via button press whether the object was newly learned or well-known. In filler trials, two different non-corresponding object fragments were randomly assigned to each object per participant.

Requiring imagery only in 25% of the trials was meant to discourage participants from initiating imagery already upon seeing the object fragment. Task preparation was rendered ineffective by the filler trials, in which invalid cues were shown. Response button assignments in the familiarity and mental imagery tasks were counterbalanced across participants. Trial types were presented in random order with short breaks after every 30 trials. Minimal knowledge, in-depth knowledge, and well-known object conditions were evenly distributed across tasks. At the end of the session, prototypical eye movements and blinks were recorded in a calibration procedure for ocular artifact correction.

EEG recording

The EEG was recorded from 56 Ag/AgCl electrodes placed according to the extended 10-20 system, initially referenced to the left mastoid. The vertical electrooculogram (EOG) was recorded from electrodes FP1 and IO1. The horizontal EOG was recorded from electrodes F9 and F10. Electrode impedance was kept below 5 kΩ. A band pass filter with 0.032 - 70 Hz, and a 50 Hz notch filter were applied; sampling rate was 250 Hz. Offline, the EEG was recalculated to average reference and low-pass filtered at 30 Hz. Eye movement and blink artifacts were removed with a spatio-temporal dipole modeling using BESA [69], based on the recorded prototypical eye movements and blinks. Trials with remaining artifacts and missing responses were discarded. The continuous EEG was segmented into epochs of 1.2 s locked to the stimulus of the main task (object picture or empty blue frame), including a 200 ms pre-stimulus baseline.

Experimental Design and Statistical Analysis

Statistical analyses were performed with R (Version 3.6.1. [70]) and the Fieldtrip toolbox [71] for Matlab (Version 2016a). Trials with unsuccessful visual search or with reaction times (RTs) shorter than 150 ms or longer than 3 *SD*s from individual participant's means were excluded from all analyses. In addition, trials with incorrect familiarity classification in the perception task were excluded from RT and ERP analyses. RTs were log transformed to approximate a normal distribution. Using the lme4 package (Version 1.1–21 [72]), accuracy and imagery success were analyzed with binomial generalized linear mixed models (GLMMs); RTs and ERPs were analyzed with linear mixed models (LMMs) [73]. LMM analyses included random intercepts and (if supported) random slopes for subjects and object identity, allowing for better generalization of results from the particular sample of participants and the set of object pictures used here. *P*-values were computed using the lmerTest package [74]. We applied sliding difference contrasts that compare mean differences between adjacent factor levels. When indicated, we reduced models by excluding non-significant interaction terms. Model selection was performed using the anova function of the stats package in R. Along with the results of the X^2-Test, we compared fit indices, Akaike information criterion (AIC) and Bayesian information criterion (BIC), that are smaller for better model fit considering the number of parameters in each model. Behavioral data were analyzed using a nested model with the factor knowledge (well-known, in-depth and minimal) nested within task (Imagery and Perception).

To address knowledge effects on ERPs during imagery and perception, we tested a priori hypotheses based on previous literature, that is, reduced P1 and enhanced N400 amplitudes with semantic knowledge in pre-specified regions of interest (ROIs). For the analysis of P1 amplitude, we averaged amplitudes within 120 to 170 ms at PO7, PO3, PO4, and PO8 (Pratt, 2011). The N400 was quantified as the mean amplitude between 300 and 500 ms at PO7, PO3, PO4, PO8, O1, Oz and O2 [24,27]. Single trial amplitudes aggregated within

ROIs and time windows were submitted to LMMs with the factors visual condition (perception, imagery and incomplete imagery) and knowledge (well-known, in-depth and minimal) as fixed effects. We fitted random structures by omitting random slopes of experimental conditions that explained zero variance, as determined by singular value decomposition.

To track the time course of activation that specifically supports imagery, we compared trials with attempted but incomplete imagery and trials with successful imagery. We also compared imagery and perception directly. To this end, we calculated each participant's average ERP in the perception, successful imagery, and incomplete imagery condition across all scalp electrodes in time windows from 0 to 540 ms. Group-level statistics were based on paired-samples *t*-tests and corrected for multiple comparisons using cluster-based permutation tests (CBPT) across time and electrodes. The cluster forming threshold was set to $p = .05$. We report differences with corrected *p*-values < .025 as statistically significant.

Based on the hypothesis that imagery might be supported in particular by configural visual processing, we looked at the N1 component. N1-amplitudes were compared in a posterior ROI consisting of PO7, PO3, PO4, PO8, O1, Oz, and O2 [52]. To adjust for latency shifts (see Results), different time windows were used for the N1 component in perception and imagery, centered around the grand mean peak latencies: For perception, we aggregated over 170 – 210 ms and for imagery (both successful and incomplete) we aggregated over 210 – 250 ms. Frontal activity that coincided with the posterior N1 was analyzed in a ROI [56] consisting of electrodes Fp1, Fpz, Fp2, AF3, AFz, AF4, F3, Fz, F4, FC1, FC2. Note that the ERP pattern at frontal sites is the opposite of that at posterior sites, therefore, we observe a frontal P1 coinciding with the posterior N1.

References

1 Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C. & de Lange, F. P. Shared representations for working memory and mental imagery in early visual cortex. *Curr. Biol.* **23**, 1427-1431, doi:https://doi.org/10.1016/j.cub.2013.05.065 (2013).

2 Kosslyn, S. M. Mental images and the Brain. *Cogn Neuropsychol* **22**, 333-347, doi:10.1080/02643290442000130 (2005).

3 O'Craven, K. M. & Kanwisher, N. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* **12**, 1013-1023, doi:https://doi.org/10.1162/08989290051137549 (2000).

4 Stokes, M., Thompson, R., Cusack, R. & Duncan, J. Top-down activation of shape-specific population codes in visual cortex during mental imagery. *J. Neurosci.* **29**, 1565-1572, doi:https://doi.org/10.1523/JNEUROSCI.4657-08.2009 (2009).

5 Cichy, R. M., Heinzle, J. & Haynes, J.-D. Imagery and Perception Share Cortical Representations of Content and Location. *Cerebral Cortex* **22**, 372-380, doi:10.1093/cercor/bhr106 (2012).

6 Dijkstra, N., Bosch, S. E. & van Gerven, M. A. J. Shared Neural Mechanisms of Visual Perception and Imagery. *Trends in Cognitive Sciences* **23**, 423-434, doi:https://doi.org/10.1016/j.tics.2019.02.004 (2019).

7 Pearson, J. The human imagination: the cognitive neuroscience of visual mental imagery. *Nat. Rev. Neurosci.* **10**, 624-634, doi:10.1038/s41583-019-0202-9 (2019).

8 Kosslyn, S. M. & Thompson, W. L. When Is Early Visual Cortex Activated during Visual Mental Imagery? *Psychol Bull* **129**, 723-746, doi:10.1037/0033-2909.129.5.723 (2003).

9 Dijkstra, N., Bosch, S. E. & van Gerven, M. A. J. Vividness of Visual Imagery Depends on the Neural Overlap with Perception in Visual Areas. *J. Neurosci.* **37**, 1367-1373 (2017).

10 Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181-204, doi:10.1017/S0140525x12000477 (2013).

11 Hohwy, J. *The predictive mind*. (Oxford University Press, 2013).

12 Friston, K. The free-energy principle: a unified brain theory? *Nature Reviews: Neuroscience* **11**, 127-138, doi:http://www.nature.com/nrn/journal/v11/n2/suppinfo/nrn2787_S1.html (2010).

13 Friston, K. The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences* **13**, 293-301, doi:10.1016/j.tics.2009.04.005 (2009).

14 Seth, A. K. in *Open MIND* (eds Thomas K. Metzinger & Jennifer M. Windt) Ch. 35(T), (MIND Group, 2015).

15 Lee, T. S. & Mumford, D. Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A: Optics and Image Science* **20**, 1434-1448 (2003).

16 Foxe, J. J. & Simpson, G. V. Flow of activation from V1 to frontal cortex in humans. *Exp. Brain Res.* **142**, 139-150, doi:10.1007/s00221-001-0906-7 (2002).

17 Friston, K. A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **360**, 815-836, doi:10.1098/rstb.2005.1622 (2005).

18 Muckli, L. & Petro, L. S. Network interactions: non-geniculate input to V1. *Current Opinion in Neurobiology* **23**, 195-201, doi:https://doi.org/10.1016/j.conb.2013.01.020 (2013).

19 Gilbert, C. D. & Li, W. Top-down influences on visual processing. *Nature Reviews: Neuroscience* **14**, 350-363, doi:10.1038/Nrn3476 (2013).

20 Bar, M. *et al.* Top-down facilitation of visual recognition. *P Natl Acad Sci USA* **103**, 449-454, doi:10.1073/pnas.0507062103 (2006).

21 Gandolfo, M. & Downing, P. E. Causal Evidence for Expression of Perceptual Expectations in Category-Selective Extrastriate Regions. *Curr Biol* **29**, 2496-2500.e2493, doi:https://doi.org/10.1016/j.cub.2019.06.024 (2019).

22 Bar, M. Visual objects in context. *Nature Reviews: Neuroscience* **5**, 617-629, doi:10.1038/Nrn1476 (2004).

23 Press, C. & Yon, D. Perceptual Prediction: Rapidly Making Sense of a Noisy World. *Curr Biol* **29**, R751-R753, doi:https://doi.org/10.1016/j.cub.2019.06.054 (2019).

24 Abdel Rahman, R. & Sommer, W. Seeing what we know and understand: How knowledge shapes perception. *Psychon Bull Rev* **15**, 1055-1063, doi:10.3758/Pbr.15.6.1055 (2008).

25 Maier, M. & Abdel Rahman, R. No matter how: Top-down effects of verbal and semantic category knowledge on early visual perception. *Cogn. Affect. Behav. Neurosci.*, doi:https://doi.org/10.3758/s13415-018-00679-8 (2019).

26 Maier, M. & Abdel Rahman, R. Native Language Promotes Access to Visual Consciousness. *Psychol. Sci.* **29**, 1757-1772, doi:10.1177/0956797618782181 (2018).

27 Rabovsky, M., Sommer, W. & Abdel Rahman, R. Depth of Conceptual Knowledge Modulates Visual Processes during Word Reading. *Journal of Cognitive Neuroscience* **24**, 990-1005 (2012).

28 Boutonnet, B. & Lupyan, G. Words Jump-Start Vision: A Label Advantage in Object Recognition. *J. Neurosci.* **35**, 9329-9335, doi:10.1523/jneurosci.5111-14.2015 (2015).

29 Collins, J. A. & Olson, I. R. Knowledge is power: How conceptual knowledge transforms visual cognition. *Psychon Bull Rev*, doi:10.3758/s13423-013-0564-3 (2014).

30 de Lange, F. P., Heilbron, M. & Kok, P. How Do Expectations Shape Perception? *Trends in Cognitive Sciences* **22**, 764-779, doi:https://doi.org/10.1016/j.tics.2018.06.002 (2018).

31 Samaha, J., Boutonnet, B., Postle, B. R. & Lupyan, G. Effects of meaningfulness on perception: Alpha-band oscillations carry perceptual expectations and influence early visual responses. *Scientific Reports* **8**, 6606, doi:10.1038/s41598-018-25093-5 (2018).

32 Lamme, V. A. Challenges for theories of consciousness: seeing or knowing, the missing ingredient and how to deal with panpsychism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **373**, 20170344 (2018).

33 Roelfsema, P. R. & Lange, F. P. d. Early Visual Cortex as a Multiscale Cognitive Blackboard. *Annu. Rev. Vis. Sci.* **2**, 131-151, doi:10.1146/annurev-vision-111815-114443 (2016).

34 Dentico, D. *et al.* Reversal of cortical information flow during visual imagery as compared to visual perception. *NeuroImage* **100**, 237-243, doi:https://doi.org/10.1016/j.neuroimage.2014.05.081 (2014).

35 Dijkstra, N., Mostert, P., de Lange, F. P., Bosch, S. & van Gerven, M. A. Differential temporal dynamics during visual imagery and perception. *Elife* **7**, e33904 (2018).

36 Lee, S.-H., Kravitz, D. J. & Baker, C. I. Disentangling visual imagery and perception of real-world objects. *Neuroimage* **59**, 4064-4073 (2012).

37 Farah, M. J., Peronnet, F., Gonon, M. A. & Giard, M. H. Electrophysiological evidence for a shared representational medium for visual images and visual percepts. *J Exp Psychol Gen* **117**, 248-257, doi:10.1037/0096-3445.117.3.248 (1988).

38 Ganis, G. & Schendan, H. E. Visual mental imagery and perception produce opposite adaptation effects on early brain potentials. *NeuroImage* **42**, 1714-1727, doi:http://dx.doi.org/10.1016/j.neuroimage.2008.07.004 (2008).

39 Suess, F. & Abdel Rahman, R. Mental imagery of emotions: Electrophysiological evidence. *Neuroimage* **114**, 147-157, doi:10.1016/j.neuroimage.2015.03.063 (2015).

40 Rossion, B. & Jacques, C. in *The Oxford handbook of event-related potential components* (eds S. J. Luck & Emily S. Kappenman) 115-142 (Oxford University Press, 2011).

41 Costa, T. L. *et al.* Neural Signatures of the Configural Superiority Effect and Fundamental Emergent Features in Human Vision. *Scientific Reports* **8**, 13954, doi:10.1038/s41598-018-32289-2 (2018).

42 Tanaka, J. W. & Curran, T. A neural basis for expert object recognition. *Psychol. Sci.* **12**, 43-47, doi:10.1111/1467-9280.00308 (2001).

43 Busey, T. A. & Vanderkolk, J. R. Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research* **45**, 431-448, doi:https://doi.org/10.1016/j.visres.2004.08.021 (2005).

44 Milivojevic, B., Corballis, M. C. & Hamm, J. P. Orientation sensitivity of the N1 evoked by letters and digits. *Journal of Vision* **8**, 11-11, doi:10.1167/8.10.11 (2008).

45 Maurer, D., Grand, R. L. & Mondloch, C. J. The many faces of configural processing. *Trends in Cognitive Sciences* **6**, 255-260, doi:https://doi.org/10.1016/S1364-6613(02)01903-4 (2002).

46 Weller, P. D., Rabovsky, M. & Abdel Rahman, R. Semantic Knowledge Enhances Conscious Awareness of Visual Objects. *Journal of Cognitive Neuroscience* **31**, 1216-1226, doi:10.1162/jocn_a_01404 (2019).

47 Maier, M., Glage, P., Hohlfeld, A. & Abdel Rahman, R. Does the semantic content of verbal categories influence categorical perception? An ERP study. *Brain Cogn.* **91**, 1-10, doi:http://dx.doi.org/10.1016/j.bandc.2014.07.008 (2014).

48 Forder, L., He, X. & Franklin, A. Colour categories are reflected in sensory stages of colour perception when stimulus issues are resolved. *Plos One* **12**, 1-16 (2017).

49 Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B. & Kuipers, J. R. Unconscious effects of language-specific terminology on preattentive color perception. *P Natl Acad Sci USA* **106**, 4567-4570, doi:10.1073/pnas.0811155106 (2009).

50 Clark, V. P., Fan, S. & Hillyard, S. A. Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Hum Brain Mapp* **2**, 170-187, doi:10.1002/hbm.460020306 (1994).

51 Di Russo, F., Martinez, A., Sereno, M. I., Pitzalis, S. & Hillyard, S. A. Cortical sources of the early components of the visual evoked potential. *Hum. Brain Mapp.* **15**, 95-111, doi:10.1002/Hbm.10010 (2002).

52 Pratt, H. in *The Oxford handbook of event-related potential components* (eds S. J. Luck & Emily S. Kappenman) 89-114 (Oxford University Press, 2011).

53 Tobimatsu, S. & Celesia, G. G. Studies of human visual pathophysiology with visual evoked potentials. *Clinical Neurophysiology* **117**, 1414-1433, doi:https://doi.org/10.1016/j.clinph.2006.01.004 (2006).

54 Kutas, M. & Federmeier, K. D. Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annu Rev Psychol* **62**, 621-647, doi:10.1146/annurev.psych.093008.131123 (2011).

55 Dijkstra, N., Zeidman, P., Ondobaka, S., van Gerven, M. A. J. & Friston, K. Distinct Top-down and Bottom-up Brain Connectivity During Visual Perception and Imagery. *Sci. Rep.* **7**, 5677, doi:10.1038/s41598-017-05888-8 (2017).

56 Gazzaley, A. *et al.* Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *P Natl Acad Sci USA* **105**, 13122-13126, doi:10.1073/pnas.0806074105 (2008).

57 Alday, P. M. How much baseline correction do we need in ERP research? Extended GLM model can replace baseline correction while lifting its limits. *Psychophysiology* **56**, e13451, doi:10.1111/psyp.13451 (2019).

58 Farah, M. J. Psychophysical evidence for a shared representational medium for mental images and percepts. *Journal of Experimental Psychology: General* **114**, 91-103, doi:10.1037/0096-3445.114.1.91 (1985).

59 Rabovsky, M., Hansen, S. S. & McClelland, J. L. Modelling the N400 brain potential as change in a probabilistic representation of meaning. *Nature Human Behaviour* **2**, 693-705, doi:10.1038/s41562-018-0406-4 (2018).

60 Sergent, C., Baillet, S. & Dehaene, S. Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience* **8**, 1391-1400, doi:http://www.nature.com/neuro/journal/v8/n10/suppinfo/nn1549_S1.html (2005).

61 Dehaene, S. & Changeux, J.-P. Experimental and theoretical approaches to conscious processing. *Neuron* **70**, 200-227 (2011).

62 Firestone, C. & Scholl, B. J. Cognition does not affect perception: Evaluating the evidence for ‘top-down’effects. *Behav Brain Sci* **39**, 1-77 (2016).

63 Lupyan, G. Objective Effects of Knowledge on Visual Perception. *Journal of Experimental Psychology: Human Perception and Performance* **43**, 794-806, doi:10.1037/xhp0000343 (2017).

64 Lupyan, G. Changing What You See by Changing What You Know: The Role of Attention. *Frontiers in Psychology* **8**, doi:10.3389/fpsyg.2017.00553 (2017).

65 Fazekas, P. & Nanay, B. Pre-Cueing Effects: Attention or Mental Imagery? *Frontiers in Psychology* **8**, doi:10.3389/fpsyg.2017.00222 (2017).

66 Emrich, S. M., Al-Aidroos, N., Pratt, J. & Ferber, S. Visual Search Elicits the Electrophysiological Marker of Visual Working Memory. *Plos One* **4**, e8042, doi:10.1371/journal.pone.0008042 (2009).

67 Woodman, G. F. & Arita, J. T. Direct Electrophysiological Measurement of Attentional Templates in Visual Working Memory. *Psychological Science* **22**, 212-215, doi:10.1177/0956797610395395 (2010).

68 Luria, R. & Vogel, E. K. Visual Search Demands Dictate Reliance on Working Memory Storage. *J Neurosci* **31**, 6199-6207, doi:10.1523/jneurosci.6453-10.2011 (2011).

69 Ille, N., Berg, P. & Scherg, M. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of Clinical Neurophysiology* **19**, 113-124 (2002).

70 R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing (2019).

71 Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience* **2011**, 1-9, doi:10.1155/2011/156869 (2011).

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

73 Frömer, R., Maier, M. & Abdel Rahman, R. Group-Level EEG-Processing Pipeline for Flexible Single Trial-Based Analyses Including Linear Mixed Models. *Frontiers in Neuroscience* **12**, 1-15, doi:10.3389/fnins.2018.00048 (2018).

74 lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-33. (2016).