Mental simulation of four visual object properties

Published in:
Journal of Cognitive Psychology

Publication status and date:
E-pub ahead of print: 04/01/2017

DOI (link to publisher):
10.1080/20445911.2017.1281283

Document Version
Publisher's PDF, also known as Version of record

Citation for the published version (APA):
de Koning, B., Wassenburg, S. I., Bos, L. T., & van der Schoot, M. (2017). Mental simulation of four visual object properties: Similarities and differences as assessed by the sentence–picture verification task. *Journal of Cognitive Psychology*, *29*(4), 420-432. Advance online publication. https://doi.org/10.1080/20445911.2017.1281283

Link to publication on the EUR Research Information Portal

Journal of Cognitive Psychology

ISSN: 2044-5911 (Print) 2044-592X (Online) Journal homepage: www.tandfonline.com/journals/pecp21

Mental simulation of four visual object properties: similarities and differences as assessed by the sentence–picture verification task

Björn B. de Koning, Stephanie I. Wassenburg, Lisanne T. Bos & Menno van der Schoot

JOURNAL OF COGNITIVE PSYCHOLOGY, 2017
VOL. 29, NO. 4, 420–432
http://dx.doi.org/10.1080/20445911.2017.1281283

Mental simulation of four visual object properties: similarities and differences as assessed by the sentence–picture verification task

Björn B. de Koning[a,b], Stephanie I. Wassenburg[a,b], Lisanne T. Bos[a] and Menno van der Schoot[a]

[a]Department of Educational Neuroscience and LEARN! Research Institute for Learning and Education, Faculty of Psychology and Education, VU University Amsterdam, Amsterdam, The Netherlands; [b]Department of Psychology, Education, and Child Studies, Erasmus School of Behavioral and Social Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands

ABSTRACT

In the sentence–picture verification (SPV) task, people read sentences implying the shape/size/colour/orientation of objects. They then verify whether pictured objects, which either *match* or *mismatch* the implied visual information mentioned in the sentence. Faster verification times on matching trials (match advantage) are considered supportive to the notion that readers perform mental simulations during sentence comprehension. This study advances this work by applying a within-subjects design to the SPV-task, enabling us to directly address the strength of and correlation between the match advantages for the properties shape, size, colour, and orientation. Results showed varying match advantages with colour showing the strongest effect, and no match advantage for orientation. Shape, size, and colour were significantly correlated, whereas there were no significant correlations with orientation. These findings suggest that interpretations of match advantages could benefit from a re-evaluation of mental simulation accountsby distinguishing between intrinsic (shape, size, and colour) and extrinsic (orientation) object properties.

ARTICLE HISTORY
Received 9 April 2015
Accepted 4 January 2017

KEYWORDS
Mental simulation; sentencepicture verification task; reading comprehension; embodied cognition

In capturing the meaning of text, readers mentally simulate the described situations and events through the reactivation of previously acquired real-world perceptual, motor, and affective experiences (Barsalou, 1999). Retrieving such an experience entails the re-enactment of the neural states from the brain systems that govern perception, action, and emotion which were recruited—and stored—at the moment the original experience was acquired. Integrating these traces of earlier experiences from multiple perceptual and motor modalities into a mental simulation enables readers to experience the described events as if they are actually part of it (for an overview, see de Koning & van der Schoot, 2013). It follows that, in such an embodied view to language comprehension, the representations involved in understanding sentences are of the same kind as the representations involved in having actual sensory and motor experiences. At present, there is a growing number of studies substantiating the claim that readers indeed activate sensory and motor information in language comprehension (Barsalou, 2008). Moreover, research has shown that action and perception circuits in the brain, contributing to comprehension, are interdependent (Pulvermüller & Fadiga, 2010). Although a coherent mental representation of text typically consists of information from all sensory modalities (Sadoski & Paivio, 2007), the visual modality in particular has incited much empirical work (de Koning & van der Schoot, 2013). In this article, we will concentrate on the work that, using the sentence–picture verification (SPV) task, has argued that visual information related to an object's properties—such as shape, size, colour, and orientation—is activated during sentence comprehension (de Koning, Wassenburg, Bos, & van der Schoot, 2016; Zwaan & Pecher, 2012; for an overview, see Horchak, Giger, Cabral, & Pochwatko, 2014).

In the SPV task, participants read sentences implying a visual property of an object like "The ranger saw the eagle in the nest" (e.g. Zwaan, Stanfield, & Yaxley, 2002). They then indicate whether or

CONTACT Björn B. de Koning b.b.dekoning@fsw.eur.nl, b.b.de.koning@vu.nl

not an object, which is depicted in such a way that it matches (i.e. an eagle with folded wings) or mismatches (i.e. an eagle with its wings outstretched) the visual property (here: object shape) as implied by the sentence, was mentioned in the sentence. Importantly, simply verifying whether the depicted object is mentioned in the sentence does not necessarily require mental simulation processes or even sentence context. Nevertheless, across numerous studies, readers appear to verify pictures faster when they match rather than mismatch the perceptual information from the preceding sentence. This so-called match advantage[1] is consistent with the idea that readers reactivate previously stored perceptual information from the brain during the processing of sentences and have this information available when verifying subsequently presented pictures, resulting in shorter verification latencies. Since the first study using the SPV task to study mental simulation of visual object properties which targeted object orientation (Stanfield & Zwaan, 2001), researchers have mainly focused on the generalisability of these results extending the findings to other visual properties such as shape (Zwaan, Stanfield, & Yaxley, 2002), visibility (Yaxley & Zwaan, 2007), colour (Connell, 2007), number (Patson, George, & Warren, 2014), distance (Vukovic & Williams, 2014), and size (de Koning et al., 2016) of described objects. Conceivably, conceptual replications of the match advantage are typically considered supportive of the general tenet of embodied approaches to language comprehension that mental representations formed during sentence comprehension contain perceptual information and that all visual aspects are grounded in the visual modality (Barsalou, 1999; Wilson, 2012).

Whilst the findings derived from all of the SPV task studies so far seem straightforward, zooming in on these findings shows a less clear and less consistent pattern of results. Whereas these variations could be actually existent representing real differences in mental simulation between different visual object properties, they could also just be the result of differences in characteristics of the SPV task used in different studies. For example, studies differed in whether participants were presented with line drawings (e.g. Stanfield & Zwaan, 2001) or photographs (e.g. Engelen, Bouwmeester, de Bruin, & Zwaan, 2011) of objects. In addition, the pictures used in prior studies inevitably varied in visual appearance such as their style and colourisation. Also, these studies varied in the populations that were studied (primary school children, Engelen et al., 2011; undergraduate psychology students, Zwaan et al., 2002; broad age range of adults, Zwaan & Pecher, 2012), the settings in which they were tested (e.g. online platforms, Zwaan & Pecher, 2012; lab cubicles, Zwaan et al., 2002; school classrooms, Engelen et al., 2011; de Koning et al., 2016), and their data trimming procedures (see Zwaan & Pecher, 2012). Moreover, although embodied accounts of language do not necessarily expect a common underlying mechanism for all kinds of mental simulations, it is often assumed that "if comprehenders perform mental simulations [for one visual dimension], then other visual dimensions should also be simulated" (Zwaan & Pecher, 2012, p. 2). This implies that readers routinely simulate all visual properties during comprehension of sentences, independently of, for example, their relevance to object identification, which is an important part of this particular task. At present, however, this assumption has only received circumstantial support given that a direct comparison between visual object properties has yet to be made. Stronger support for this assumption would be obtained if prior results on each of these individual visual object properties could be replicated in such a direct comparison.

Based on these considerations, the present study employed, as far as we know for the first time, a within-subjects design to the SPV task so as to investigate the mental simulation of the visual object properties shape, size, colour, and orientation during sentence comprehension. The within-subjects design enables us to address at least two aspects related to the mental simulation of these visual object properties that research using the SPV task so far has not been able to answer. First, it offers the possibility to take a look at the fact that across studies employing the SPV task, the magnitude of the match advantage appears to vary from one visual property to the other. Second, it enables an investigation of the stability of an individual's reliance on mental simulations across different visual object properties.

Regarding the first aspect, varying match advantage magnitudes could be interpreted as differences

[1]Note that this does not necessarily imply that participants' responses are facilitated in the match condition rather than slowed down in the mismatch condition. With the term "match advantage"—as opposed to mismatch advantage—we simply indicate faster reaction times for matching than for mismatching trials.

in the extent to which visual object properties are mentally simulated in the visual modality. For example, properties that are more relevant to object identification than others may be stronger simulated than properties that are less central to object identification. Furthermore, the degree to which a visual property contributes to a representation may change as a function of context (van Dam, van Dijk, Bekkering, & Rueschemeyer, 2012). As such, if the process of mental simulation actually differs across different visual object properties, this would challenge the claim—made by some strong embodied cognition proponents—that lexically driven semantic representations are unaffected by context (e.g. Boulenger, Hauk, & Pulvermüller, 2009; Schuil, Smits, & Zwaan, 2013). So far, a relatively large and robust match advantage (averaged over studies approximately 63 ms, see Zwaan & Pecher, 2012) is typically reported for studies involving the mental simulation of object shape (e.g. Rommers, Meyer, & Huettig, 2013; Zwaan et al., 2002; for a more elaborated discussion, see Zwaan & Pecher, 2012). Recently, de Koning et al. (2016) have reported a match advantage for the mental simulation of object size which was of comparable magnitude (61 ms for children and 69 ms for adults). Also for object colour large match advantages have been reported (averaged over studies approximately 121 ms, see Zwaan & Pecher, 2012), although it could be questioned to what extent the colour effect is robust given that the opposite pattern of results, a mismatch advantage, has also been observed (Connell, 2007). Nevertheless, Zwaan and Pecher (2012) have argued that future studies are more likely to show match advantages for object colour rather than mismatch advantages. For object orientation, however, much lower match advantages are reported across studies (38 ms, see Zwaan & Pecher, 2012). Importantly, there are even studies that fail to find a match advantage for orientation at all (e.g. Rommers, Meyer, & Huettig, 2013).

Although these findings suggest clear differences in the reported match advantages for each of the visual object properties, to the best of our knowledge, no study has yet attempted to actually test this assumption directly. Studies have typically focused on just a single visual object property at the time (e.g. Connell, 2007; Zwaan et al., 2002), have intermixed different visual object properties (i.e. shape and orientation) in one SPV task (e.g. Engelen et al., 2011), or if more than one visual object property was investigated in a study have treated them as between-subject factors in different SPV tasks (Zwaan & Pecher, 2012). Hence, empirical evidence is yet inconclusive about the relative magnitude of the match advantages of different visual object properties observed in previous studies.

Interestingly, embodied accounts of language comprehension do not seem to make any predictions that would lead one to hypothesise differential match advantages depending on the visual property (i.e. shape, size, colour, orientation) that is studied. An approach that could provide some more guidance in this respect is to consider the kind of visual object property. Shape, size, and colour are intrinsic visual properties that distinguish an object from another. These properties are independent of relations to other objects or a frame of reference, and thus are invariant. Instead, orientation is an extrinsic visual property in that it is not inherent to an object, but rather varies depending on aspects like the observer (e.g. rotating or tilting his/her head), the particular condition of observation (e.g. looking straight ahead whilst standing upright or when lying on your side), or the way an object is positioned (e.g. holding a sword upright or directing it at someone in front of you) (Borghi, Caligiore, & Scorolli, 2010; Scorolli & Borghi, 2008). It is important to realise that such variations in the orientation of an object do not change the object itself, but only how the object is held or perceived.

Applying this to the previously reported match advantage findings, it could be argued that match advantages vary depending on whether sentences contain an intrinsic (shape, size, and colour) or extrinsic (orientation) visual object property. Therefore, the present study examines the extent to which the distinction between extrinsic and intrinsic properties is useful in helping us to explain the observed differences in the magnitude of the match advantages. More precisely, it examines whether, as suggested by previous literature (e.g. Rommers et al., 2013; Zwaan & Pecher, 2012), the match advantage is smaller for extrinsic properties (i.e. orientation) as compared to intrinsic properties (i.e. shape, size, and colour). This advances prior empirical and theoretical work on the SPV task providing a more detailed perspective on whether the nature and qualities of a visual property influence reader's mental simulation of visual object properties during sentence comprehension.

The second aim of this study is to investigate the stability of an individual's reliance on visual mental simulations across the four visual object properties.

Assuming that an individual consistently engages in mental simulation irrespective of which visual object property is involved (e.g. Vukovic & Williams, 2014; Wassenburg & Zwaan, 2010; Yaxley & Zwaan, 2007), we could expect to find correlated match advantages across all visual object properties. This would be consistent with the idea that different visual object properties are mentally simulated in the visual modality (Barsalou, 1999). If, however, match advantages do not appear to be correlated across object properties, readers may not engage in visual mental simulations during reading as consistently as was previously thought. For example, based on the intrinsic–extrinsic distinction the orientation of an object is not a dominant feature in visually defining or identifying an object and its affordances (Harris & Dux, 2005). As such, when verifying a depicted object of which only the orientation mismatches the orientation implied by the preceding sentence in the SPV task, the reader may not experience a semantic mismatch. In contrast, for intrinsic visual object properties verifying a matching (e.g. picture of a white bear following the sentence "He saw a bear at the North pole") or mismatching (e.g. picture of a brown bear following the sentence "He saw a bear at the North pole") picture involves a semantically different object. Whereas the matching and mismatching pictures visually only differ in colour, they are associated with a (partly) different semantic network (for a similar suggestion, see Zwaan & Pecher, 2012).

So, a specific sentence context (e.g. describing a protagonist who manually manipulates an object) may cause orientation to be processed differently than the visually dominant intrinsic properties shape, size, and colour (Hoenig, Sim, Bochev, Herrnberger, & Kiefer, 2008). If mental simulations are sensitive to a context that changes the relevance of modality-specific properties, this could potentially influence which brain areas will be activated during mental simulation. Depending on sentence context, mental simulation of orientation may, for example, involve the activation of motor-related areas rather than mental simulation in the visual modality only. That this possibility of different processing is not unlikely, becomes clear when we move beyond the SPV task and consider the literature in adjacent fields of research. For example, by now several studies have shown that for object manipulation the grasp component of a movement is influenced by intrinsic object properties, whereas the reaching component of a movement is influenced by extrinsic object properties, suggesting that different processes are associated with intrinsic and extrinsic object properties (Gentilucci, Chieffi, Scarpa, & Castiello, 1992; Jeannerod, 1981; Jeannerod, Arbib, Rizzolatti, & Sakata, 1995). The within-subjects design in the SPV task employed in the present study provides a first step to investigate whether or not an individual stably relies on context-independent mental simulation processes across visual object properties.

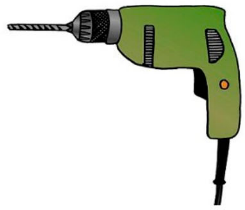
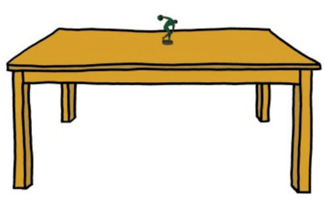
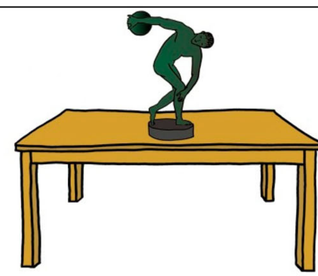
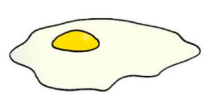
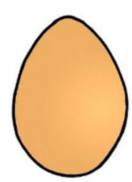
Taken together, the present study aimed to extend previous SPV task research by taking a within-subjects design approach to the SPV task so as to provide insight into two yet unresolved issues regarding the mental simulation of visual object properties during sentence comprehension. In this study, participants thus completed a SPV task four times, one for each of the visual object properties shape, size, colour, and orientation. Our first interest was to investigate the suggestion that the magnitude of the match advantages differs depending on the visual object property involved, suggesting that mental simulation may be more dependent upon the nature and qualities of the visual object property than previously thought. Based on previous empirical work, we may expect intrinsic visual object properties (i.e. shape, size, and colour) to show significantly stronger match advantages than extrinsic visual object properties (i.e. orientation). A second aim of this study was to examine whether the match advantages for the four visual object properties would show a correlation, and thus be indicative of a common mental process to stably simulate visual object properties.

Method

Participants

The participants were 169 university students recruited from different educational programmes of a large Dutch university. Consistent with biased gender ratios in these programmes, there were 125 women and 44 men, with an average age of 20.97 years ($SD = 2.19$ years). The participants provided written consent indicating that they voluntarily took part in the experiment. Participants received course credit for their participation. All participants participated in a counterbalanced within-subjects design with the factors Visual Property, consisting of four levels (i.e. shape, size, colour, and orientation), and Matching, consisting of two levels

Table 1. Examples of sentence–picture pairs (match and mismatch) for each visual object property.

Visual property	Sentence-picture combinations
Orientation	  The handyman put the drill against the wall / ceiling (De klusser zette de boormachine tegen de muur / het plafond)
Color	  The girl licked the ice cream with chocolate / vanilla flavor (Het meisje likte aan het ijsje met chocoladesmaak / vanillesmaak)
Size	  The woman saw the sculpture in the windowsill / garden (De vrouw zag het beeld op de vensterbank / in de tuin)
Shape	  The chef took the egg out of the skillet / box (De kok haalde het ei uit de koekenpan / het doosje)

Note. Dutch sentences are given in parentheses. The underlined words imply one of two versions of the visual property; the first option matches the left picture (and thus mismatches the right picture) and the second option matches the right picture.

(i.e. match vs. mismatch), both of which are discussed in more detail below.

Materials

SPV task

In the SPV task, readers read a sentence implying a visual property of an object (e.g. shape and orientation) and have to indicate whether or not the object shown in a subsequently presented picture was mentioned in the sentence. In this study, four SPV tasks were used, each targeting a different visual property for an object: shape, size, colour, and orientation. These tasks were based on the original experimental sentence and picture stimuli that have been used to study the object properties shape, size, colour, and orientation, developed by Stanfield and Zwaan (2001), de Koning et al. (2016), Connell (2007), and Zwaan et al. (2002), respectively. Where appropriate, we translated items into Dutch, rephrased sentences, and constructed new items (e.g. the colour task developed by Connell (2007) contained only half as many items as the other SPV tasks). An example of each, as used in the current study, is provided in Table 1.

In all four SPV tasks, a set of experimental sentences (presented in Dutch) and pictures were used. Sentences were of the format: subject [The cook], verb [saw], noun (critical object) [the egg], prepositional phrase [in the skillet]. The prepositional phrase was critical for implying a certain shape, size, colour, or orientation of the object. For the experimental sentences, there were always two versions that only differed in their prepositional phase, and thus implying different versions of a visual property of the object. For example, the sentence "The man saw the eagle in the sky" suggests a different shape of the critical object than the sentence "The man saw the eagle in the nest". Participants read one of the two versions of each sentence and subsequently saw a picture that either matched or mismatched the visual property implied in the sentence (see Table 1). All pictures were coloured drawings made by a qualified draftsman. This ensured that all pictures were comparable regarding style (e.g. thickness of lines, brightness, and colour use). All pictures were scaled to occupy a square of approximately 15 × 15 cm on the computer screen. Importantly, each picture only showed the critical object, except for the pictures in the SPV task involving implied object size where the critical object was always presented on the same table (see Table 1). As determining object size requires contextual cues, this table provided a necessary reference point which enabled participants to "read off" an object's size directly (de Koning et al., 2016). The SPV tasks involving the visual object properties shape, colour, and orientation contained 24 experimental sentence–picture items, whereas for the SPV task involving implied size there were 28 experimental sentence–picture items.

For each SPV task, we created four experimental lists by crossing the two versions of experimental sentences and the two versions of pictures, using a Latin Square design. As a result, each list included one of the four possible sentence–picture combinations for an object. Each participant saw one of four lists per SPV task that he or she completed. Across the four lists, all item combinations were used equally often. On each list, half (i.e. 12) of the experimental sentence–picture items matched whereas the other half (i.e. 12) mismatched on the visual property of interest. As the answer to the 24 experimental items always required a "yes" response, an equal number of filler items were added to balance responses. Filler items consisted of a sentence which had the same amount of words as the experimental sentences which was followed by a picture unrelated to the critical object mentioned in the sentence. Filler items thus always required a "no" response.

Procedure

All four SPV tasks were administered individually in a single session in a quiet cubicle in the university's lab facility. The order of presentation of the SPV tasks was counterbalanced across the participants to diminish unwanted serial order effects (e.g. fatigue, improved performance due to experience with the task). Participants were instructed to read each sentence at their own pace and to indicate as fast and accurate as possible whether or not the depicted object had been mentioned in the sentence. For the visual property size, participants were told that objects were shown on a table and they were instructed to pay attention to the depicted object on the table. Upon the first encounter of each SPV task, participants were presented with a few practice trials. Next, the experimental and filler trials were presented in a random order. Each trial started with a horizontally and vertically centred sentence

Table 2. Picture verification latencies and accuracy.

	Percentage correct	
	Match	Mismatch
Shape	98 (4)	96 (7)
Size	95 (6)	93 (8)
Colour	97 (5)	91 (11)
Orientation	98 (4)	98 (5)

Note: Standard deviations are given in parentheses.

on the computer screen, displayed in a black 24-point Courier New Bold font against a white background. Participants pressed the spacebar when they had understood the sentence, after which a 500 ms fixation cross appeared, followed by a picture. Participants indicated whether or not the pictured object was mentioned in the preceding sentence by pressing the keys on the keyboard marked by a green sticker ("yes" response) and a red sticker ("no" response). The whole experiment lasted approximately 40 minutes.

Results

Data trimming for the four SPV tasks consisted of removing all reaction times (RTs) faster than 300 ms and slower than 3000 ms (Connell, 2007), and all responses that were more than 2.5 standard deviations from the participant's mean RT in the relevant condition. This resulted in removal of less than 5% of the data in all SPV tasks (shape: 3.20%; size: 4.18%; colour: 4.14%; and orientation: 2.52%).[2] In line with the procedure used by Zwaan and Pecher (2012), for every SPV task separately data of participants with an accuracy score of .80 or lower were left out of our analyses as such unusually low accuracy scores on this task are considered to not accurately reflect the cognitive processes the task aims to target. Data of 47 participants who performed below the 80% accuracy criterion were excluded. Because some participants performed below this criterion on more than one SPV task the total number of excluded participants per SPV task (i.e. shape: $n = 9$; size: $n = 20$; colour: $n = 30$; orientation: $n = 2$) does not correspond to the number of all excluded participants. Of the original 169, there were 122 participants who scored more than 80% accurate on all four tasks. Presumably, participants got tired after performing more than one SPV task resulting in low accuracy scores on subsequent SPV tasks. For further mean RT analyses, incorrect responses were eliminated. Table 2 shows the mean accuracy scores for the four visual properties in the match and mismatch conditions. Importantly, the accuracy scores as shown in Table 2 did not show evidence for a speed-accuracy trade-off when compared to the mean RTs for matching and mismatching trials (see Figure 1). That is, faster response times for matching versus mismatching items did not coincide with reduced accuracy on those items. Therefore, comparison and interpretation of these RTs are warranted.

First of all, we aimed to test whether the match advantage in the SPV task differs between visual properties (i.e. shape, size, colour, and orientation). Therefore, we conducted a 4 (Visual Property: shape, size, colour, and orientation) × 2 (Matching: match vs. mismatch) repeated measures ANOVA. Bayesian analyses were performed in addition to regular null hypothesis significance testing using the statistical software JASP (2016). The posterior probabilities (BF_{01})[3] favouring the alternative hypothesis are reported. If BF_{01} is smaller than 1, adding the effect to the model makes the evidence against the null hypothesis more compelling. The results of our analyses indicated that there was a significant main effect of Visual Property, $F(3,363) = 67.45$, $p < .001$, $\eta_p^2 = .36$, $BF_{01} < 0.001$, indicating that overall RTs differed between the four tasks, with the slowest RTs for colour and size and faster RTs for shape and orientation (see Figure 1). Furthermore, a significant main effect for Matching indicates that overall participants responded faster to matching sentence–picture trials than to mismatching sentence–picture trials, $F(1,121) = 26.57$, $p < .001$, $\eta_p^2 = .18$, $BF_{01} = 0.013$. This effect, however, differed as a function of the visual property implied in the SPV task. This can be observed in Figure 1, and was evidenced by a significant interaction effect between Visual Property and Matching, $F(3,363) = 10.01$, $p < .001$, $\eta_p^2 = .08$, $BF_{01} = 0.27$. *Post hoc* pairwise comparisons with Bonferroni correction (resulting in a one-tailed alpha at .012), were performed to

[2]Given that in some of the original studies on the match advantage the analyses were performed on median RT scores (shape: Zwaan et al., 2002 and orientation: Stanfield & Zwaan, 2001), we have also conducted analyses on median RT scores. These analyses yielded the same statistical pattern as the results of the analysis on the mean RTs. Therefore, only results for mean RTs are reported.

[3]The BF_{01} is an odds ratio for the null/alternative hypotheses given the data, where values larger than 1 indicate more evidence for the null hypothesis, and values smaller than 1 indicate more evidence for the alternative hypothesis (Masson, 2011). More specifically, a BF_{01} smaller than 0.10 can be interpreted as strong evidence (Jarosz & Wiley, 2014).

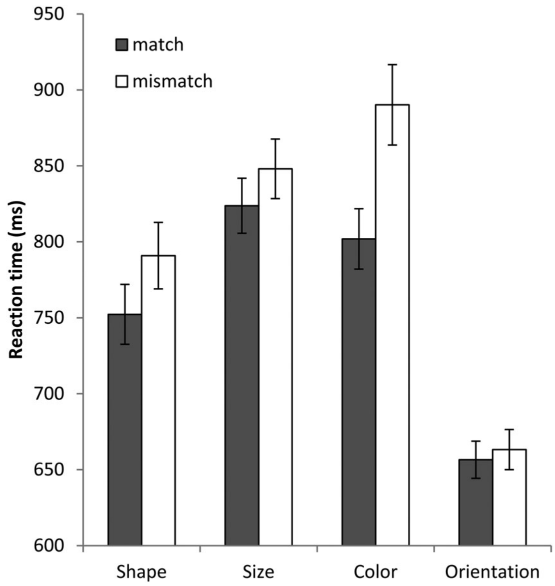


Figure 1. Average RTs in milliseconds for the four visual object properties of the SPV task in the match and mismatch conditions (error bars depict standard errors of the mean).

investigate the match advantage (i.e. shorter RTs for matching than for mismatching trials) in the four visual properties separately. The match advantage was largest for the visual property colour, $\Delta RT = 88$ ms, $t(138) = 5.69$, $p < .001$, $d = .48$, $BF_{01} < 0.001$, followed by shape, $\Delta RT = 39$ ms, $t(159) = 3.42$, $p < .001$, $d = .27$, $BF_{01} = 0.022$, size, $\Delta RT = 24$ ms, $t(148) = 2.31$, $p = .011$, $d = .07$, $BF_{01} = 0.43$, and orientation, $\Delta RT = 7$ ms, $t(166) = 0.92$, $p = .181$, $d = .07$, $BF_{01} = 4.702$. Together, these results show small to large match advantages for the intrinsic visual properties shape, size, and colour, and no match advantage for the extrinsic visual property orientation.

To provide additional evidence of such a distinction, a *post hoc* contrast analysis was performed comparing intrinsic visual properties (i.e. shape, size, and colour) to extrinsic visual properties (i.e. orientation). The results showed a significant interaction effect between Visual Property (intrinsic vs. extrinsic) and Matching (match vs. mismatch),[4] $F(1, 129) = 21.16$, $p < .001$, $\eta_p^2 = .14$, indicating that the effect of Matching was significantly larger for the three SPV tasks involving intrinsic visual properties than for the SPV task involving extrinsic visual properties. These findings suggest that obtaining a match advantage in the SPV task is influenced by, among other things, the nature (i.e. intrinsic vs. extrinsic) of the visual object property.

Table 3. Correlations between the difference scores of the four visual object properties.

	1.	2.	3.	4.
1. Shape		.38* * (<0.001)	.18* (0.568)	.01 (9.173)
2. Size			.25* * (0.098)	.09 (3.117)
3. Colour				−.01 (9.962)
4. Orientation				

Note: BF_{01} is given in parentheses.
**$p < .01$.
*$p < .05$.

A second issue we wanted to investigate is the relation between match advantages for the visual properties shape, colour, orientation, and size. If an individual stably relies on mental simulation to represent visual object properties in reading comprehension, the performance on the four SPV tasks implying these visual object properties are expected to be correlated. Match–mismatch difference scores for the four visual properties were calculated for each participant. Based on Cook's distance (>.10), 8 outlying difference scores were removed (1% of the data). Visual analysis of the scatterplots did not reveal any other outliers. A correlation analysis was conducted on the difference scores of the four SPV tasks. As shown in Table 3, the correlations between the three visual object properties colour, shape, and size were significant. We found no significant correlations, however, between the object property orientation and the other visual object properties. So, consistent with the RT analyses, there was a significant correlation among the intrinsic visual properties and these did not correlate with the only visual object property that is extrinsic in nature. Taken together, these results challenge the hypothesis of a common underlying mechanism involved in perceptually simulating the visual properties of objects in the SPV task.

Discussion

This study applied a within-subjects design to the SPV task in order to address two yet unresolved issues regarding the mental simulation of the visual object properties shape, size, colour, and orientation. First, we directly investigated whether, as suggested by prior research, the match advantages for these visual object properties actually show a distinct pattern of findings. We used the

[4]This type of contrast cannot be performed by JASP yet and was therefore tested using SPSS.

distinction between intrinsic (i.e. shape, size, and colour) and extrinsic (i.e. orientation) visual properties as a potential source for possible variations. Second, to investigate whether all four visual properties share a common underlying mechanism on which readers stably rely, we examined whether there is a correlation among the match advantages of the four visual object properties.

The results of this study show that some visual object properties elicit stronger match advantages than others. This confirms our presumption derived from prior (between-subject) SPV task studies (e.g. de Koning et al., 2016; Stanfield et al., 2002; Zwaan & Pecher, 2012), and indicates that differences in the match advantages of the four visual object properties exist independent of variations in stimuli, settings, and populations. In fact, we used a set of stimuli that was carefully constructed so that the stimuli were as comparable as possible. For example, all pictures were drawn by a professional draftsman to ensure that across SPV tasks there were no differences in terms of the pictures' appearance (e.g. style, brightness, colour use). Together, these results are compatible with the suggestion that match advantage differences are most likely due to the extent to which readers engage in visual mental simulation.

In particular, our findings demonstrate the expected—variations in the—match advantages for the visual object properties shape, size, and colour. Of these, colour showed the strongest match advantage, which is consistent with the findings previously reported by Zwaan and Pecher (2012). Using a larger set of colour stimuli than prior studies, our study now puts the balance on three strong match advantages against one mismatch advantage for colour (Connell, 2007). This indicates that the colour effect does not appear as weak and unstable as initially thought (also see Zwaan & Pecher, 2012). Shape and size showed comparable match advantages, with shape exhibiting a slightly stronger effect. This finding supports de Koning et al.'s (2016) suggestion that a close correspondence exists between these two visual object properties. Interestingly, this pattern showed up even though overall participants took longer to verify the pictures on the size task. Our findings suggest that it does not matter for obtaining a match advantage that an extra object (i.e. a table on which the critical object was displayed to "read off" objet size directly) was presented; it only seemed to take somewhat longer to obtain such an effect as two objects rather than one had to be dealt with when verifying the pictures. Whilst these findings together are in line with our expectations, the match advantages obtained in our study are somewhat smaller than those in previous studies. One possible reason for this is that, in contrast to previous SPV task studies, the present study employed a within-subjects design in which each participant completed four SPV tasks each targeting another visual object property instead of having to complete just one SPV task. Consistent with this, Zwaan and Pecher (2012) reported that match advantages did not disappear as participants completed multiple SPV tasks (in a between-subject design), but it did reduce the strength of the match advantage.

Importantly, our findings thus demonstrate a match advantage, of varying strength, for all three intrinsic visual object properties. Instead, the only extrinsic visual object property investigated in this study, orientation, did not show such a match advantage. This finding may not be very surprising considering that findings on the orientation effect appear weak and not very consistent (e.g. Zwaan, 2014). Whereas some researchers (e.g. Zwaan & Pecher, 2012) have found a match advantage, albeit a small one compared to other visual object properties, others (e.g. Rommers et al., 2013) have failed to obtain a match advantage for orientation. Such divergent findings are typically explained by arguing that failures to replicate the orientation effect are due to slight deviations from Zwaan et al.'s (2001) original orientation experiment and thus do not allow for meaningful comparisons (Zwaan, 2014). However, the fact that our study contained four SPV tasks which all followed exactly the same procedures and were comparable in the design of the stimuli enables us to make straightforward comparisons between match advantages on the SPV tasks. So, it does not seem likely that the failure to find an orientation effect in our study emanates from differences between SPV tasks. Rather, the present findings more likely fit the idea put forward in the Introduction that the match advantage is sensitive to the kind of visual object property (i.e. intrinsic vs. extrinsic) that should be mentally simulated when reading a sentence. This suggests that mental simulations are more dependent upon the nature and qualities of visual object properties than previously thought. Apparently, visual properties that are more dominant in defining an object, which is of particular importance for identifying and verifying depicted objects in the SPV task, are stronger mentally simulated in the visual modality than non-dominant visual properties.

The findings following from the correlation analyses are in line with this interpretation. Specifically, significant correlations were found between all of the three intrinsic visual object properties (i.e. shape, size, and colour); the extrinsic visual object property orientation appeared not to be significantly correlated with any of the intrinsic visual object properties. So, the results suggest that a reader who mentally simulates the shape implied in a sentence (i.e. showing a match advantage) also stably relies on mental simulation of the size and colour implied in a sentence, but does not necessarily, or at least to a lesser extent, mentally simulate an object's orientation as implied in a sentence. We interpret this finding as challenging the theoretical stance adopted in previous SPV task studies, which assumes that mental simulation in a visual modality represents the core common underlying mechanism involved in processing both the intrinsic and extrinsic visual object properties investigated in this study (e.g. Horchak et al., 2014; Zwaan & Pecher, 2012). Together, our findings suggest that it could advance the interpretation of match advantages if distinctions were made between the effects obtained for intrinsic visual object properties and those obtained for extrinsic ones. Following up on this, we suspect readers to engage in visual mental simulations, when reading a sentence implying an object's shape, size, or colour (i.e. intrinsic object properties). In this view (but see the limitations for an alternative interpretation), match advantages are believed to arise from a relatively larger overlap between activated brain patterns resulting from (re)enacting perceptual information of the described situation and seeing the object picture (Barsalou, 2008). We assume that, because intrinsic object properties are important to defining a particular object, for matching trials—part of—the semantic network that is activated by an object with the particular shape/size/colour implied in the sentence appears to be consistent with that needed to accurately verify the picture. Rather, for mismatching trials at least part of the semantic network that is activated does not overlap with the information that is activated by seeing the picture. Having this information available may interfere with the verification process, resulting in longer verification latencies in the mismatch condition. Presumably, these processes together contribute to the match advantage for intrinsic visual object properties. In contrast, because orientation is less important in defining a particular object, the semantic network that is activated by a vertical object does not necessarily differ from the semantic network activated by the horizontal version of the object. As such, readers may not experience a conceptual mismatch in the visual representation.

Importantly, based on our study, there is reason to assume that extrinsic visual object properties (i.e. orientation) implied in a sentence relies less, if at all, on activation of the visual modality. As described in the Introduction, extrinsic properties are dependent on the observer, the particular condition of observation, or the way an object is positioned. We speculatively attribute our finding that the extrinsic property orientation is the only visual object property that "falls out" to the fact that readers, depending on the sentence context, might recruit other modalities to which orientation is more relevant (e.g. the orientation of an object may be more relevant for reaching for the object than for visually identifying it). This is in accordance with research suggesting that the modality of representation recruited to perform mental simulations is dependent on the dynamic relationship between the individual, movement, and environment (Stevens, 2005Stevens, 2005). In addition, these results are compatible with the idea that the mental simulation process is not purely automatic, but is—at least to some extent—flexible (i.e. activation of our perceptual systems can be present or absent) and dependent on context (van Dam, van Dijk, Bekkering, & Rueschemeyer, 2012). In other words, it is possible that mental simulation of visual object properties is modulated by the relevance of modality-specific properties in a given context.

Alternatively, it could still be argued that extrinsic visual object properties initially elicit automatic activations in visual areas (Yaxley & Zwaan, 2007), but that the corresponding mismatches are quickly nullified by relying on additional processes such as mental rotation. In this case, to make a verification decision, readers try to map the presented picture to the visual-spatial mental image they created from the object described in the sentence. Depending on the extent to which the implied and pictured orientation of an object match, readers (have to) engage in a 90° mental rotation of the object so as to let it correspond to the presented picture. The time needed to be able to make a decision about whether the orientation in the sentence and picture match likely changes accordingly. This explanation, however, seems less likely because a previous study has shown no reliable correlation between mental rotation and the magnitude of the orientation match advantage (Stanfield & Zwaan, 2001). Furthermore, an additional time-consuming mental rotation component does not fit with the fact that in our study overall RTs were faster for orientation than for the other visual properties. Whether the above suggestions adequately account for the match advantage differences found between intrinsic and extrinsic visual object properties in the SPV task remains to be examined in future research, for example by studies that more directly address the recruited brain areas with time-sensitive measurements (e.g. MEG/EEG) or potential additional processing mechanisms involved in this task.

Although by applying a within-subjects design our study provides some novel perspectives on the processes that are elicited by verifying pictures after reading sentences, the present study has some limitations that should be mentioned. First, it should be acknowledged that the results of the present study cannot exclude a non-embodied account of reading comprehension. For example, reading a sentence implying visual object properties and then verifying a picture does not necessarily require that "visual" representations are generated during reading. During or after reading, participants may simply have accessed linguistic and world knowledge and generate verbal expectations about the described situation (e.g. bears at the North Pole tend to be white), which may not necessarily derive from perceptual experience of seeing white bears (but for example from reading about bears). Upon seeing a picture, participants could label the bear as a polar bear and then make faster verification responses if the label they generated from the picture contextually matches the sentence interpretation. To what extent reduced or no reliance on perceptual experiences/representations influences processing of one or more intrinsic/extrinsic visual object properties needs to be investigated in future studies. Second, even though we have tried to make all four SPV tasks as comparable as possible (i.e. controlling for test setting, visual appearance of stimuli, data trimming methods, etc.) to minimalise external influences and to directly address the strength of and correlation between the match advantages for the visual object properties shape, size, colour, and orientation, the comparison between visual object properties, however, remains between items. We cannot exclude the possibility that differences between the four visual properties in the magnitude of the match advantage are influenced by the particular (differences between) items that we have used. Additionally, the saliency of object properties between SPV tasks may potentially confound the results in this study. Possibly, shape, size, and colour were more salient in the depiction of the object than orientation and a less salient property may be more easily integrated in a sentence representation even if it mismatches. Therefore, future studies should control for potential differences in saliency. Third, employing a within-subject design with four SPV tasks may have caused participants to get tired or confused resulting in lower accuracy. Compared to previous SPV studies, a relatively a large group of participants scored below the 80% accuracy cut off score on one of four tasks, with the highest mistake rate for the colour task. To maintain a high accuracy rating for all tasks, future studies may consider not assessing all tasks in one session.

In conclusion, this study demonstrates that, consistent with our predictions, different kinds of visual object properties (intrinsic/extrinsic) have somewhat different effects in a SPV task. The three intrinsic visual object properties shape, size, and colour all showed clear match advantages, with colour showing the strongest effect, followed by shape and size which exhibited more or less comparable effects. The extrinsic visual object properties orientation, however, did not show a significant match advantage. Moreover, we found significant correlations among the intrinsic visual object properties and these did not correlate with the extrinsic visual property orientation. These findings suggest that the distinction between intrinsic and extrinsic visual object properties might be helpful to explain the variations in the reported match advantages and provides a useful supplement to the mental simulation account. Importantly, more research is required to substantiate this claim and to clarify what mechanisms are responsible for the fact that some visual object properties (shape, size, and colour) are strongly activated in the visual modality during language comprehension, whilst others (orientation) are not. This study provides a starting point from which such endeavours can be further explored.

Acknowledgments

The authors would like to thank the master students for helping with the data collection and Eline Bos for drawing the pictures. We also wish to thank Louise Connell for sharing the materials of the colour task, based on which we constructed the colour task stimuli for the present study.

Funding

This research was supported by a grant from the Netherlands Organization for Scientific Research/Dutch Program Council for Educational Research (NWO/PROO) for the project "Learning to read with imagination" (#411-11-807).

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Stephanie I. Wassenburg http://orcid.org/0000-0002-9007-3136

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660. doi:10.1017/S0140525X99002149

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. doi:10.1146/annurev.psych.59.103006.093639

Borghi, A. M., Caligiore, D., & Scorolli, C. (2010). Objects, words, and actions. Some reasons why embodied models are badly needed in cognitive psychology. In V. Capecchi & M. Buscema (Eds.). *Applications of mathematics in models, artificial neural networks and arts. Mathematics and society* (pp. 99–112). Berlin: Springer.

Boulenger, V., Hauk, O., & Pulvermüller, F. (2009). Grasping ideas with the motor system: Semantic somatotopy in idiom comprehension. *Cerebral Cortex*, *19*(8), 1905–1914. doi:10.1093/cercor/bhn217

Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, *102*, 476–485. doi:10.1016/j.cognition.2006.02.009

de Koning, B. B., & van der Schoot, M. (2013). Becoming part of the story! Refueling the interest in visualization strategies for reading comprehension. *Educational Psychology Review*, *25*, 261–287. doi:10.1007/s10648-013-9222-6

de Koning, B. B., Wassenburg, S. I., Bos, L. T., & van der Schoot, M. (2016). Size does matter: The implied object size is represented during language comprehension. *Discourse Processes*, article online in advance of print. Retrieved from http://dx.doi.org/10.1080/0163853X.2015.1119604

Engelen, J. A. A., Bouwmeester, S., de Bruin, A. B. H., & Zwaan, R. A. (2011). Perceptual simulation in developing language comprehension. *Journal of Experimental Child Psychology*, *110*, 659–675. doi:10.1016/j.jecp.2011.06.009

Gentilucci, M., Chieffi, S., Scarpa, M., & Castiello, U. (1992). Temporal coupling between transport and grasp components during prehension movements: Effects of visual perturbation. *Behavioural Brain Research*, *47*, 71–82. doi:10.1016/S0166-4328(05)80253-0

Harris, I. M., & Dux, P. E. (2005). Orientation-invariant object recognition: Evidence from repetition blindness. *Cognition*, *95*(1), 73–93. doi:10.1016/j.cognition.2004.02.006

Hoenig, K., Sim, E., Bochev, V., Herrnberger, B., & Kiefer, M. (2008). Conceptual flexibility in the human brain: Dynamic recruitment of semantic maps from visual, motor, and motion-related areas. *Journal of Cognitive Neuroscience*, *20*(10), 1799–1814. doi:10.1162/jocn.2008.20123

Horchak, O. V., Giger, J., Cabral, M., & Pochwatko, G. (2014). From demonstration to theory in embodied language comprehension: A review. *Cognitive Systems Research*, *29–30*, 66–85. doi:10.1016/j.cogsys.2013.09.002

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, *7*, 2–9. doi:10.7771/1932-6246.1167

JASP. (2016). Retrieved from http://jasp-stats.org/

Jeannerod, M., Arbib, M. A., Rizzolatti, G., & Sakata, H. (1995). Grasping objects: The cortical mechanisms of visuomotor transformation. *Trends in Neurosciences*, *18*, 314–320. doi:10.1016/0166-2236(95)93921-J

Jeannerod, M. (1981). Intersegmental coordination during reaching at natural visual objects. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 153–168). Hillsdale, NJ: Erlbaum.

Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*(3), 679–690. doi:10.3758/s13428-010-0049-5

Patson, N. D., George, G., & Warren, T. (2014). The conceptual representation of number. *The Quarterly Journal of Experimental Psychology*, *67*(7), 1349–1365. doi:10.1080/17470218.2013.863372

Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, *11*(5), 351–360. doi:10.1038/nrn2811

Rommers, J., Meyer, A. S., & Huettig, F. (2013). Object shape and orientation do not routinely influence performance during language processing. *Psychological Science*, *24*, 2218–2225. doi:10.1177/0956797613490746

Rommers, J., Meyer, A. S., & Huettig, F. (2013). Object Shape and Orientation Do Not Routinely Influence Performance During Language Processing. *Psychological Science*, *24*(11), 2218–2225.

Sadoski, M., & Paivio, A. (2007). Toward a unified theory of reading. *Scientific Studies of Reading*, *11*(4), 337–356. doi:10.1080/10888430701530714

Schuil, K. D., Smits, M., & Zwaan, R. A. (2013). Sentential context modulates the involvement of the motor cortex in action language processing: An fMRI study. *Frontiers in Human Neuroscience*, *7*, 100. doi:10.3389/fnhum.2013.00100

Scorolli, C., & Borghi, A. M. (2008). Language and embodiment. *Anthropology and Philosophy*, *9*(1–2), 7–23.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*, 153–156. doi:10.1111/1467-9280.00326

Stevens, J. A. (2005). Interference effects demonstrate distinct roles for visual and motor imagery during the mental representation of human action. *Cognition*, *95* (3), 329–350.

van Dam, W. O., van Dijk, M., Bekkering, H., & Rueschemeyer, S. A. (2012). Flexibility in embodied lexical-semantic representations. *Human Brain Mapping*, *33*, 2322–2333. doi:10.1002/hbm.21365

Vukovic, N., & Williams, J. N. (2014). Automatic perceptual simulation of first language meanings during second language sentence processing in bilinguals. *Acta Psychologica*, *145*, 98–103. doi:10.1016/j.actpsy.2013.11.002

Wassenburg, S. I., & Zwaan, R. A. (2010). Readers routinely represent implied object rotation: The role of visual experience. *The Quarterly Journal of Experimental Psychology*, *63*(9), 1665–1670. doi:10.1080/17470218.2010.502579

Wilson, M. (2012). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. doi:10.3758/BF03196322

Yaxley, R. H., & Zwaan, R. A. (2007). Simulating visibility during language comprehension. *Cognition*, *105*(1), 229–236. doi:10.1016/j.cognition.2006.09.003

Zwaan, R. A. (2014). Replications should be performed with power and precision: A response to Rommers, Meyer, and Huettig (2013). *Psychological Science*, *25*, 305–307. doi:10.1177/0956797613509634

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PLoS One*, *7*(12). Retrieved from http://dx.doi.org/10.1371/journal.pone.0051382

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