Journal of Experimental Child Psychology 110 (2011) 659–675

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Perceptual simulation in developing language comprehension

Jan A.A. Engelen [a,*], Samantha Bouwmeester [a], Anique B.H. de Bruin [a,b], Rolf A. Zwaan [a]

[a] Institute of Psychology, Erasmus University Rotterdam, 3000 DR Rotterdam, The Netherlands
[b] Department of Educational Development and Research, Maastricht University, 6200 MD Maastricht, The Netherlands

ARTICLE INFO

Article history:
Received 11 April 2011
Revised 21 June 2011
Available online 30 July 2011

Keywords:
Language comprehension
Embodied cognition
Perceptual simulation
Language development
Word reading
Reading comprehension

ABSTRACT

We tested an embodied account of language proposing that comprehenders create perceptual simulations of the events they hear and read about. In Experiment 1, children (ages 7–13 years) performed a picture verification task. Each picture was preceded by a prerecorded spoken sentence describing an entity whose shape or orientation matched or mismatched the depicted object. Responses were faster for matching pictures, suggesting that participants had formed perceptual-like situation models of the sentences. The advantage for matching pictures did not increase with age. Experiment 2 extended these findings to the domain of written language. Participants (ages 7–10 years) of high and low word reading ability verified pictures after reading sentences aloud. The results suggest that even when reading is effortful, children construct a perceptual simulation of the described events. We propose that perceptual simulation plays a more central role in developing language comprehension than was previously thought.

Introduction

Consider the sentences “The ranger saw the eagle in the sky” and “The ranger saw the eagle in the nest.” The former refers to an eagle with outstretched wings, whereas the latter refers to an eagle that has its wings drawn in. Theories of embodied language comprehension predict that when readers process these sentences, their mental representation of the eagle changes accordingly (Zwaan, Stanfield, & Yaxley, 2002). To test this prediction, researchers have used the perceptual mismatch

* Corresponding author.
E-mail address: engelen@fsw.eur.nl (J.A.A. Engelen).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2011.06.009

paradigm (Stanfield & Zwaan, 2001), in which a participant reads or listens to a sentence describing a situation and is subsequently presented with a picture of an object involved in that situation. Critically, the shape or orientation of the depicted object either matches or mismatches the shape or orientation implied by the description. For example, "The ranger saw the eagle in the sky" can be followed by a picture of a flying eagle (match) or a perched eagle (mismatch). It has been found that participants need more time to verify or name mismatching pictures relative to matching pictures (Dijkstra, Yaxley, Madden, & Zwaan, 2004; Hirschfeld & Zwitserlood, 2010; Holt & Beilock, 2006; Kaup, Yaxley, Madden, Zwaan, & Lüdtke, 2007; Madden & Dijkstra, 2010; Madden & Zwaan, 2006; Stanfield & Zwaan, 2001; Zwaan et al., 2002). A symbolic theory of language comprehension, in which the eagle might be represented as a list of features or a node in a propositional network, cannot easily account for this mismatch effect. Instead, the mismatch effect suggests that meaning is instantiated by the partial reactivation and integration of previous perceptual experiences.

Although the studies listed above typically involved undergraduates or (in some cases) older adults, it is as of yet unclear whether children construct perceptual simulations during language comprehension. On the one hand, work such as Mandler's (1992, 2010) highlights the role of spatial representations in the development of conceptual knowledge. From there, it is a small step to proposing that perceptual simulations are central to young children's comprehension processes. On the other hand, the extent to which perceptual simulations are constructed as a function of language comprehension is constrained by domain expertise (e.g., Holt & Beilock, 2006) and processing capacity (e.g., Madden & Zwaan, 2006). These limitations may be especially restricting when children learn to read. This is because during the early stages of reading, processing resources are primarily allocated to breaking the orthographic code (Perfetti, 1985) rather than to mapping this code onto meaningful representations. Therefore, we may expect the use of perceptual simulations to follow different developmental trajectories for reading and listening. The current study explores this issue in 7- to 13-year-olds. This approach potentially informs theories of language development by showing how children represent meaning under different linguistic modalities and, at the same time, feeds into theories of embodied cognition by showing to what extent developmental factors constrain the utility of perceptual simulations. The following sections discuss in more detail the task of language comprehension and the role of domain expertise and processing efficiency.

Situation models

It is a generally held view that when people comprehend language, they create a mental representation of the described state of affairs rather than of the text itself (Johnson-Laird, 1983; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). This representation is referred to as a *situation model.* According to theories of embodied language comprehension, perceptual–motor simulations—not amodal propositions—are the building blocks of situation models. Readers and listeners construct these simulations by reactivating and integrating traces of previous experience distributed across multiple perceptual and motor modalities in the brain (Barsalou, 1999; Zwaan & Madden, 2005). The recruitment of perceptual and motor systems has been demonstrated in a host of behavioral tasks, showing that comprehenders simulate not only the implied shape and orientation of objects but also many other perceptual features of a situation such as an object's direction of motion (Glenberg & Kaschak, 2002; Kaschak et al., 2005; Zwaan, Madden, Yaxley, & Aveyard, 2004; Zwaan & Taylor, 2006), the rate and length of fictive motion (Matlock, 2004), the axis along which an action takes place (Richardson, Spivey, McRae, & Barsalou, 2003), the part of the visual field where a scene takes place (Bergen, Lindsay, Matlock, & Narayanan, 2007), the visibility of objects through an obscuring medium (Yaxley & Zwaan, 2007), and the sounds produced by entities (Brunyé, Ditman, Mahoney, Walters, & Taylor, 2010). Efforts to build a situation model for larger stretches of discourse also engage visual processing mechanisms. For instance, memorizing a dot matrix for a later recognition test, which involves a visual–spatial load, has a stronger degrading effect on comprehension of short texts than remembering a letter string, which involves a verbal load (Fincher-Kiefer, 2001).

The mounting evidence for the automatic activation of experiential traces during language comprehension does not imply that these are the only means that individuals have at their disposal for representing meaning. One can achieve at least a rudimentary understanding of discourse by making use of the co-occurrences of linguistic forms. Latent semantic analysis (LSA) (Landauer & Dumais, 1997) uses an algorithm that maps words into a high-dimensional semantic space. Linguistic forms (e.g., words, phrases, whole texts) are compared in this space, resulting in a cosine value representing their semantic relatedness. LSA does a remarkable job at simulating human performance, for instance, by correctly answering multiple-choice textbook questions after being trained on the content of the textbook (Landauer, Foltz, & Laham, 1998). Nevertheless, as long as linguistic forms, which are inherently abstract, merely refer to other linguistic forms, which are also inherently abstract, the semantic relatedness values are essentially meaningless. For a linguistic form to be meaningful, it must be grounded in experience outside the network (Glenberg & Robertson, 2000; Harnad, 1990; Searle, 1980). These concerns notwithstanding, linguistic representations may be functional to on-line language comprehension. This might work as follows. According to principles of content addressability and encoding specificity, the information in memory that is most similar to the cue becomes active most rapidly (e.g., Tulving & Thomson, 1973). When an incoming word is recognized, activation spreads to associated linguistic representations and subsequently to perceptual representations (see also Paivio, 1986). One possible function of the activation of linguistic representations is to anticipate upcoming information (e.g., Barsalou, Santos, Simmons, & Wilson, 2008; Zwaan, 2008). For example, on hearing the word *bird*, the perceptual system is set up to activate a representation of not only *bird* but also *sky*, *fly*, and *wings*, thereby facilitating activation and integration of these concepts in case any of these words will actually follow.

Whether linguistic or perceptual–motor representations dominate an individual's response on a given task depends on the individual's domain-specific knowledge and processing capacity. Owing to the time-constrained nature of most comprehension tasks, it matters how quickly a perceptual representation is activated as a function of hearing or seeing a word and how quickly sentence context can be used to constrain the developing simulation (Madden & Zwaan, 2006). Moreover, although language affords the description and simulation of unfamiliar situations, possessing a rich network of experiential traces should facilitate the rapid activation of a trace that is appropriate in a given linguistic context. Importantly, both processing efficiency and domain-specific knowledge increase throughout childhood, but neither may be adequately developed in 7-year-olds to support the on-line construction of perceptual simulations. We turn to these issues below.

Domain expertise

In many ways, development during childhood parallels that from novice to expert in specific domains. Holt and Beilock (2006) investigated whether domain expertise had a bearing on individuals' comprehension of descriptions of domain-specific situations. Novice and expert football and hockey players performed a sentence–picture verification task in the perceptual mismatch paradigm described above. When sentences dealt with everyday scenarios or actions anyone might perform (e.g., "The woman put the umbrella in the closet"), both groups performed accurately, making the correct decision on 96% of the trials. In addition, both novice and expert athletes showed a mismatch effect (i.e., they responded more quickly to pictures that matched the previously presented sentence). However, when the sentences described sport-specific scenarios (e.g., "The coach saw the defenseman stop the kick"), both groups still performed accurately, but only expert athletes showed a mismatch effect. These findings suggest that possessing perceptual–motor representations depends on experience interacting with objects and performing the actions in question.

Analogously, adults and older children are more likely to possess rich networks of perceptual representations than younger children, even for everyday scenarios. This is reflected in the growth of the depth of vocabulary knowledge (e.g., Lahey, 1988; Ouellette, 2006). For example, a 13-year-old may have a multitude of perceptual traces associated with the word *pigeon*, having seen it in flight, walking, and perched on a roof, having studied its colors in picture books, and having heard its cooing. This allows the individual to readily use an appropriate trace for constructing a simulation of the sentence "Bob saw the pigeon in the sky." Suppose that a 7-year-old's only experience with a pigeon is seeing it perched. The child may represent the flying pigeon by second-order grounding (Harnad, 1990), for instance, by using his or her perceptual knowledge about other flying birds. In line with the findings by

Holt and Beilock (2006), this way of grounding language in experience might not be adequately efficient to support time-constrained comprehension.

Processing efficiency

Another crucial component of language comprehension is the ability to hold words and clauses in memory while processing other words and clauses until both can be integrated. This ability is measured by the reading span task (Daneman & Carpenter, 1980). Reading span is operationalized as the number of words that an individual can keep in memory while giving sensibility judgments about a set of unrelated sentences. As such, this measure directly taps into the efficiency of the comprehension process. Previous research has demonstrated that only high-span comprehenders immediately apply sentence-level context during comprehension (Madden & Zwaan, 2006; Van Petten, Weckerly, McIsaac, & Kutas, 1997). In one study, Madden and Zwaan (2006) compared high- and low-span comprehenders on their performance on a sentence–picture verification task in the perceptual mismatch paradigm. The location was stated first, so that the target object was always the last word in the sentence (e.g., "In the pot there was spaghetti"). When the picture was presented 750 ms after the offset of the target word, both low- and high-span comprehenders showed a mismatch effect. When the picture was presented immediately after the offset of the target word, only high-span comprehenders showed a mismatch effect. These findings suggest that high-span comprehenders were efficient at activating a contextually appropriate perceptual representation of the target word. Low-span comprehenders, on the other hand, were slower to construct a perceptual simulation and, therefore, relied on a linguistic representation.

Reading span increases during childhood (e.g., Case, Kurland, & Goldberg, 1982; Chiappe, Hasher, & Siegel, 2000), suggesting that older children use their available processing resources more efficiently than younger children. A possible explanation is that older children have developed stronger activation links between words and their associated perceptual representations. The reinforcement of such basic-level processes may improve the efficiency of the comprehension process as a whole (MacDonald & Christiansen, 2002). Accordingly, the ability to bring sentence context to bear on perceptual simulations may be limited at the age when children start learning to read but considerably better developed after several years of education.

To summarize, relatively inefficient processing found in children and a lack of experience with the objects and actions being talked about may pose a threshold for constructing perceptual simulations during comprehension. In the current study, children from Grades 2 to 6 (7- to 13-year-olds) performed a picture verification task in the perceptual mismatch paradigm. The choice of these age groups is nonarbitrary with regard to reading education. In Grade 2, the focus shifts from decoding to comprehension. By the end of Grade 6, formal reading instruction has typically finished, although development still continues after that. Hence, these age groups capture the stages during which the propensity to construct perceptual simulations during language comprehension is likely to change.

Experiment 1

Method

Participants

Children ($N = 140$, 62 boys and 78 girls) from Grades 2 through 6 in an ethnically heterogeneous primary school in a large urban area in The Netherlands participated in the study. Each grade contributed 28 children. Ages ranged from 7.5 to 9.3 years in Grade 2 (mean = 8.3), from 8.6 to 9.9 years in Grade 3 (mean = 9.1), from 9.5 to 11.0 years in Grade 4 (mean = 10.3), from 10.5 to 12.4 years in Grade 5 (mean = 11.2), and from 11.7 to 13.3 years in Grade 6 (mean = 12.3).

Participants were screened for abnormal comprehension by their teachers. This screening was complemented with standardized tests of comprehension on which all participating children performed at age-appropriate levels. Caretakers were informed about the research and gave passive consent before the start of the experiment.

Materials

We constructed 42 experimental sentence pairs of the format "*Agent* saw the *object* in/on the *location*" (see Appendix B). Each sentence implied a distinct shape (e.g., "Bob saw the pigeon in the nest" vs. "Bob saw the pigeon in the sky") or orientation of the same object (e.g., "John saw the nail in the wall" vs. "John saw the nail in the floor"). The only difference within a given sentence pair was on the last noun and, in a few cases, on the preposition. Some location nouns were used in two sentence pairs, but care was taken that participants saw each location noun only once. All sentences were in Dutch. The sentences were prerecorded by an adult male native speaker of Dutch and edited to terminate at the offset of the last word.

For each experimental sentence pair, we selected one picture of the described object. This picture matched the object's shape or orientation implied by the location in one of the sentences but mismatched it in the other (see Appendix A). The pictures were full-color photographs obtained from various web libraries and scaled to occupy an area of approximately 10×10 cm.

In addition to the experimental items, 76 filler items were constructed, 60 of which featured unrelated pictures and 16 of which featured a picture of the location. These served to balance the number of affirmative and negative responses and to prevent participants from merely paying attention to the object noun. The total set of stimuli, intended to be used in both experiments reported here, consisted of 118 sentence–picture pairs, 58 requiring an affirmative response and 60 requiring a negative response.

To ensure that sentences were clearly audible and did not contain any unfamiliar words, we conducted a pilot study with students from Grades 2, 3, and 6 ($N = 12$). No problems with sentences or individual words were reported. We also checked whether the pictures were unambiguous and adequately fitted the described entities. Items that were responded to incorrectly by more than 3 participants were not included in the later experiment.

The picture verification and motor speed tasks were run using E-Prime stimulus presentation software (Schneider, Eschman, & Zuccolotto, 2002). Responses were registered using a custom-made response box with two large buttons, 4 cm in diameter each, labeled *no* (left) and *yes* (right), approximately 20 cm apart from each other.

Procedure

The experiment took place in a quiet room within the school environment, where participants were seated in front of a computer screen. The experiment started with a simple button-press task that used the same setup as the picture verification task. This was because we expected substantial variance in response times due to differences in motor speed. To be able to correct for these differences, we measured the participants' reaction speed in the absence of higher order cognitive processes. In a given trial, a cross appeared on the screen either left- or right-aligned. Participants were asked to respond by pressing a button on the corresponding side of the response box. Participants rested their hands on the buttons and pressed as quickly as possible. In total, 10 trials were presented in random order at an interval of 1000 ms. Only the 5 right-aligned trials were registered because only the right hand would be used for giving correct responses in the picture verification task.

Subsequently, the experimenter explained to participants that they were about to listen to a set of sentences and that each time a picture would be shown afterward. Participants were asked to rest their hands on the buttons labeled *yes* and *no* and to press as quickly as possible after they had determined whether the depicted object had been mentioned in the sentence. They listened to the sentences and kept their eyes focused on a fixation point in the center of the screen. This point was replaced by the picture after 1000 ms following the offset of the sentence. Participants began with 10 practice trials consisting of 5 related and 5 unrelated pictures. Next, they completed a sequence of 59 trials, including 21 experimental trials. Each participant performed 10 or 11 match trials and 10 or 11 mismatch trials, all requiring an affirmative response. In addition, there were 8 fillers requiring an affirmative response and 30 fillers requiring a negative response. All trials were presented in random order. The experiment took approximately 15 min to complete.

It is important to note that both experiments in the current study took place within a single session. No separate practice trials were offered in between. The order of the experiments was counterbalanced, so that half of the participants performed Experiment 2 before entering Experiment 1 and

Table 1
Accuracy and response times for experimental trials in Experiment 1.

Grade	Condition	Accuracy (%)		Response time (ms)	
		Mean	*SD*	Mean	*SD*
2	Match	97.5	5.0	1265	518
	Mismatch	95.0	9.1	1336	555
3	Match	94.0	14.1	926	428
	Mismatch	95.0	17.5	962	425
4	Match	94.6	7.1	918	386
	Mismatch	95.7	6.8	903	380
5	Match	95.7	7.1	911	414
	Mismatch	95.4	7.3	963	453
6	Match	98.9	3.1	780	289
	Mismatch	94.6	7.3	823	356

vice versa. Items were also counterbalanced across experiments, so that participants did not see the same picture in both experiments, and each picture occurred as often in Experiment 1 as it did in Experiment 2.

Results

Preliminary analyses

Two experimental items were removed prior to the statistical analyses owing to their large number of incorrect responses. Next, all trials with response latencies smaller than 300 ms or larger than 3000 ms were removed, yielding an additional loss of less than 1% of the data. The average proportion of correct responses for all remaining trials, including fillers, was .96 (SD = .07). Filler items were not included in the reaction time analyses, but this check was performed to ensure that participants complied with the instructions and were not biased toward either affirmative or negative responses. The high percentage of correct responses indicates that participants adequately understood the procedure. The percentages of correct responses and average response times (collapsed over trials) for matching and mismatching experimental trials are given in Table 1. The percentages of correct responses were similar across conditions, indicating that mismatching trials were not more likely to elicit a negative response, warranting further comparison of response times.

Analysis of response times

The common method of analysis used in response time research is to aggregate the observations over replications per participant. A disadvantage of this aggregation is that the sample size that is used is reduced to the number of participants in the sample, thereby reducing statistical power. As an alternative, a mixed model can be used where random effects are included to model the dependencies between a participant's replications. We used multilevel analysis with items at the lowest level and participants at the highest level. The dependent variable was response time, and a random intercept was estimated to model the dependencies between trials within a participant.

To correct the effects of the variables of interest for other sources of variation in response times, we conducted a hierarchical regression analysis in which we first included the fixed and random intercept and motor speed in the model. Table 2 shows the statistical tests for the mixed model effects. The random intercept was significant, indicating that after correcting for motor speed, the variance in response times between children was larger than zero. This justifies the inclusion of the random intercept in the model. Motor speed significantly predicted response times (b = 1.43, SE = 0.12, r = .64,[1] $p < .001$). Thus, the faster the performance in the motor speed task, the faster the performance in the picture verification task.

[1] The effect size r was calculated as follows: $r = \sqrt{\frac{\left(\frac{b}{SE}\right)^2}{\left(\frac{b}{SE}\right)^2 + df}}$ (Rosnow & Rosenthal, 2005).

Table 2
Hierarchical regression analysis of response times in Experiment 1.

Block	Predictor	*Wald*	*df*	*p*
1	Intercept	1532.58	1	<.001
	Random intercept	260.97	1	<.001
	Motor speed	139.19	1	<.001
2	Condition	7.00	1	.008
	Grade	61.86	4	<.001
	Condition × Grade	4.25	4	.37

In the second step, the variables condition (match vs. mismatch), grade, and Condition × Grade were added to the model. The main effect for condition was significant ($b = 81.50$, $SE = 30.81$, $r = .23$, $p = .008$). Corrected for grade and motor speed, responses to matching pictures were approximately 81 ms faster than responses to mismatching pictures. The main effect for grade was also significant, $Wald(4) = 61.86$, $p < .001$. Post hoc comparisons with Bonferroni correction applied revealed a significant difference between Grades 2 and 3 ($b = -247.58$, $SE = 48.04$, $r = .41$, $p < .001$). The negative coefficient indicates that children in Grade 3 responded faster than children in Grade 2 even after the correction for motor speed. Contrasts between other grades did not approach significance.

Finally, the interaction effect for Condition × Grade did not approach statistical significance, $Wald(4) = 4.25$, $p = .37$. Thus, we found no evidence that the mismatch effect was influenced by grade.

Discussion

The main finding of Experiment 1 was that pictures were verified faster when they matched the preceding sentence than when they mismatched the preceding sentence. A straightforward interpretation of this finding is that participants had activated a mental representation of the target word that shared certain perceptual features with the picture probe. This primed the perceptual system with at least enough precision to speed up recognition of matching pictures relative to mismatching pictures (see also Hirschfeld & Zwitserlood, 2010). Crucially, the appropriate shape or orientation of the target word could be derived only by combining the object and the location that the nouns in the sentence referred to. For example, in the sentence "Martin saw the screw in the wall," the horizontal orientation of the screw is not stated explicitly but needs to be inferred by meshing the affordances of a screw with those of a wall. Because neither a screw nor a wall does by itself enforce a horizontal representation, the mismatch effect must stem from perceptual simulation and not from associative priming.

Grade had a substantial impact on response times, even when corrected for motor speed. Specifically, children in Grade 3 responded faster than those in Grade 2. Although we did not predict this increase in speed at this specific stage of development, it might be informative with respect to children's decision-making ability in language-based tasks. Alternatively, it is possible that the motor task did not capture all of the variance related to motor speed and that older children were faster simply due to more efficient response execution.

Importantly, although children in higher grades showed a nominally larger mismatch effect than children in lower grades, the interaction between condition and grade did not approach significance. So, whereas response latencies decreased with age, the additional time for verifying mismatching pictures remained constant. What explains this pattern? We speculate that the age-related variation in response times was distributed across recognition of the depicted object, accessing the name of that object, comparison of that name with the words in the sentence, and an affirmative or negative response based on that comparison. An alternative scenario, involving spillover processing of the sentence during the presentation of the picture probe for younger children, is less likely for two reasons. First, this would require children to either process the sentence and the picture in parallel or suppress the picture until they had constructed a complete mental representation of the sentence. Given the saliency of the pictures, it is more likely that children immediately shifted their attention to the picture, performing the verification task using whatever representation of the sentence that was available to them. Second, if continuing processing of the sentence were the case, children could have used the picture to aid their construction of a mental representation of the sentence. In that case, they would have exhibited less additional processing time for mismatching pictures.

In summary, Experiment 1 supports the notion that 7- to 13-year-olds simulate the implied shape and orientation of objects. The size of the mismatch effect did not increase as a function of grade, suggesting that even in Grade 2 language about everyday situations is grounded in experience. With the current materials, we could not detect specifications beyond shape and orientation, so the possibility that older children formed richer situation models than younger children cannot be ruled out. Even so, this does not undermine the notion that young children appear to activate and integrate perceptual memory traces while comprehending spoken language. Experiment 2 investigated whether the same holds for written language.

Experiment 2

Although there is evidence that a general cognitive mechanism underlies comprehension of spoken, written, and even nonlinguistic information (e.g., Gernsbacher, 1985), we believe that it is important to distinguish between written and spoken language, especially when reading is a newly acquired skill. Glenberg, Gutierrez, Levin, Japuntich, and Kaschak (2004) pointed out that beginning readers often fail to map the words in written text to their referents, as opposed to words in speech. There are at least two reasons for this. First, spoken language is often used in highly determined contexts. When a child is first exposed to spoken language, there is a consistent, natural, and repeated association between the words being uttered and the objects and events being referred to (Masur, 1997). For example, a caregiver may talk about "the bottle" while holding a bottle or say "wave bye-bye" while actually waving (Glenberg et al., 2004). When a child learns to read, this association is broken. Written language often deals with objects and events outside of the reader's physical environment, so their referents need to be retrieved from memory. Second, when children need to read text themselves, their attention may be chiefly directed toward getting the spelling-to-sound conversions right instead of retrieving and integrating the appropriate meaning representations (Perfetti, 1985). This allocation of resources is formalized in the reader model (Just & Carpenter, 1992), according to which readers prioritize basic processes at the cost of higher order comprehension processes. In line with this, there are numerous studies showing that less skilled word reading is associated with poor comprehension (e.g., Muter, Hulme, Snowling, & Stevenson, 2004; Perfetti & Hart, 2001; Shankweiler, 1989).

From these considerations, it follows that even if children are proficient at constructing perceptual simulations in the domain of oral language, this is not necessarily true for the domain of written language. If less skilled word readers do not reliably activate a perceptual representation as a function of seeing a word or are compromised on the processing resources needed for using the linguistic context, we expect them not to show a mismatch effect for sentences they need to read themselves. This was investigated in Experiment 2, where we compared participants of high and low word reading ability on their performance on a sentence–picture verification task.

Method

Participants

The participants in this experiment were children from Grades 2 through 4 ($N = 78$, 38 boys and 40 girls) who also participated in Experiment 1. There were 27 children from Grade 2 (mean age = 8.2 - years, range = 7.5–9.3), 28 children from Grade 3 (mean age = 9.1 years, range = 8.6–9.9), and 23 children from Grade 4 (mean age = 10.3 years, range = 9.5–11.0).

Participants were assigned to groups according to their word reading ability scores. Word reading ability was measured with the Three-Minute Test (TMT) (Verhoeven, 1995), a standardized test that takes into account both speed and accuracy. Children read words from three lists of increasing complexity (monosyllabic words with single consonants, monosyllabic words with consonant clusters, and polysyllabic words). Each list was shown for 1 min. The score was calculated by subtracting the number of incorrectly pronounced words from the number of correctly pronounced words. Less skilled readers (n = 20, mean age = 8.4 years, range = 7.6–9.4) had TMT scores of 55 or lower (M = 44, SD = 7). Skilled readers (n = 58, mean age = 9.4 years, range = 7.5–11.0) had TMT scores of 56 or higher (M = 76, SD = 12). This split corresponds to the norm of minimally acceptable word reading ability in Grade 2. As such, less skilled word readers in our sample can be assumed to be representative of less skilled word readers in the population in general.

Materials

The sentence–picture pairs were identical to those in Experiment 1 except that sentences were presented as text on the screen. The sentences were centered and displayed in a black 18-point Courier New font against a white background. To ensure that the sentences did not surpass the children's word reading ability, we conducted a pilot study with students from Grades 2, 3, and 6 (N = 12). There were no problems with the object and location nouns in the experimental sentences. However, some names were pronounced incorrectly by 1 or more children. These names were replaced along with the incorrectly pronounced object nouns in filler sentences.

Procedure

Participants sat in front of a computer screen and were instructed to read the sentences aloud. Incorrectly pronounced or skipped words were recorded by the experimenter. At the offset of the last word, the experimenter immediately pressed a button to replace the sentence by a fixation point. After 1000 ms, this point was replaced by the picture probe.

Half of the participants had not completed Experiment 1 before and started with 10 training trials. All participants then performed a sequence of 59 trials, including 21 experimental trials. The experimental trials consisted of 10 or 11 match trials and 10 or 11 mismatch trials, all requiring an affirmative response. In addition, there were 8 fillers requiring an affirmative response and 30 fillers requiring a negative response. The trials were presented in random order. The experiment took approximately 15 min to complete.

Results

Preliminary analyses

Trials that contained incorrectly pronounced or skipped words (1.4% of all trials) were removed from the data set. The two low-accuracy pictures from Experiment 1 figured in Experiment 2 as well and were also removed. Next, all trials with response times longer than 3000 ms were eliminated, yielding an additional loss of less than 1% of data. The average proportion of correct responses for all remaining trials, including fillers, was .95 (SD = .08). Filler items were not included in the reaction time analyses, but this check was performed to ensure that participants were not biased toward either affirmative or negative responses. The high percentage of correct responses indicates that participants adequately understood the procedure. The percentages of correct responses and average response times (collapsed over trials) for matching and mismatching experimental trials are given in Table 3. The percentages of correct responses were similar across conditions, indicating that mismatching trials were not more likely to elicit a negative response and warranting further comparison of response times.

Word reading and grade were significantly correlated (r = .70, p < .001). To rule out potential multicollinearity issues, we checked the unique contributions of both predictors by computing their

Table 3
Accuracy and response times in Experiment 2.

Word reading	Condition	Accuracy (%)		Response time (ms)	
		Mean	SD	Mean	SD
Low (n = 20)	Match	94.7	10.9	1286	513
	Mismatch	93.5	15.4	1375	515
High (n = 58)	Match	96.9	5.7	979	387
	Mismatch	95.4	6.4	1037	430

Table 4
Hierarchical regression analysis of response times in Experiment 2.

Block	Predictor	*Wald*	*df*	*p*
1	Intercept	1648.35	1	<.001
	Random intercept	100.85	1	<.001
	Motor speed	201.99	1	<.001
2	Condition	4.88	1	.027
	Word reading	0.44	1	.510
	Grade	27.11	2	<.001
	Condition × Word Reading	0.16	1	.690
	Condition × Grade	1.22	2	.540

partial correlations. Controlling for word reading, grade and response times were significantly correlated ($r = -.19$, $p < .001$). Controlling for grade, word reading and response times were significantly correlated ($r = -.13$, $p < .001$). We concluded that both variables accounted for unique portions of variance in response times. Hence, both variables were included in subsequent analyses.

Analyses of response times

We used multilevel analysis with items at the lowest level and participants at the highest level. As in Experiment 1, we conducted a hierarchical regression analysis in which we included the fixed and random intercept and motor speed first. Table 4 shows the statistical tests for the mixed model effects. The random intercept was significant, justifying the inclusion of a random intercept in the model. Motor speed significantly predicted response times ($b = 1.95$, $SE = 0.19$, $r = .71$, $p < .001$). Thus, the faster the response on the motor speed task, the faster the response on the picture verification task.

In a second step, the following block of predictors was entered: condition (match vs. mismatch), word reading (high vs. low), grade, Condition × Word Reading, and Condition × Grade. The main effect for condition was significant ($b = 79.98$, $SE = 36.19$, $r = .26$, $p = .027$). Correcting for grade, word reading, and motor speed, responses to matching pictures were approximately 80 ms faster than responses to mismatching pictures. The main effect for word reading did not approach significance ($b = -53.31$, $SE = 80.14$, $r = .08$, $p = .51$). This means that we found no difference between skilled and less skilled word readers in response times when controlling for grade and motor speed.[2] The main effect for grade was significant, $Wald(2) = 27.11$ $p < .001$. Post hoc comparisons with Bonferroni correction applied revealed a significant difference between Grades 2 and 3 ($b = -368.16$, $SE = 100.20$, $r = .40$, $p < .001$). The negative coefficient indicates that children in higher grades responded faster than children in lower grades, even after correcting for motor speed and word reading skill. The difference between Grades 3 and 4 did not approach statistical significance.

The interaction effect for Condition × Word Reading was not significant ($b = 28.82$, $SE = 71.10$, $r = .05$, $p = .69$). This means that word reading ability had no bearing on the size of the mismatch effect. Finally, the interaction effect for Condition × Grade did not approach significance, $Wald(2) = 1.22$, $p = .54$. Thus, there is no indication that the mismatch effect changed as a function of grade.

Discussion

In Experiment 2, a sentence–picture mismatch effect was observed for written language. The interaction between condition and word reading skill did not approach statistical significance, indicating that skilled and less skilled word readers showed a similar advantage for matching pictures. Although we anticipated that nonfluent word reading would interfere with the process of retrieving and integrating the appropriate memory traces, it did not preclude a perceptual simulation of the described situation.

[2] Alternatively, word reading scores could be treated as a continuous variable. Because using raw word reading scores did not change the pattern of results, we used two discrete groups for ease of interpretation.

It should be noted that children did not read the sentences silently; rather, they read them aloud. It has been found that the performance aspect of reading aloud in the presence of an audience, such as an experimenter, hampers comprehension relative to reading aloud to oneself or reading silently (Holmes, 1985). If anything, this should decrease the likelihood of perceptual simulation taking place, making the observed mismatch effect even more surprising.

One possible explanation for not finding larger effects related to word reading is that the words and syntactic constructions used in the sentences were too easy for individual differences in word reading ability to emerge. However, the reading times per sentence were longer for the less skilled readers (*M*s = 5.0 and 3.6 s), suggesting that their performance was well below ceiling.

Parallel to Experiment 1, a main effect was observed for grade. Older children responded faster than younger children, even when motor speed was partialed out. The decrease in response times was most notable between Grades 2 and 3. Again, it is not likely that this difference is explained by additional time needed to process the sentence alone; rather, it is the sum of the time needed to recognize the picture, access the name of the picture, compare the name with the words in the sentence, and the actual response based on that comparison. This would be consistent with the previous finding that less skilled readers show impaired performance relative to skilled readers on tasks that require explicit comparison between a test probe and the preceding context (Long, Seely, & Oppy, 1999).

General discussion

Two experiments addressed the question of whether 7- to 13-year-olds construct perceptual simulations during language comprehension. The results suggest that they do while listening to sentences in Experiment 1 and while reading sentences aloud in Experiment 2. The children's responses in a sentence–picture verification task were consistent with the hypothesis that they had formed a perceptual simulation of the described situation that they had constructed within 1000 ms after the offset of the critical location noun. Although response times were consistently longer after written sentences than after spoken sentences, the mismatch effect was comparable in magnitude across experiments and also comparable to that obtained with adults with similar procedures (e.g., Stanfield & Zwaan, 2001; Zwaan et al., 2002). Moreover, the mismatch effect emerged as a robust phenomenon in that it did not increase as a function of grade or word reading skill. This is surprising given previous demonstrations of the constraining role of expertise (Holt & Beilock, 2006) and processing capacity (Madden & Zwaan, 2006). Before discussing the implications of these findings for the development of language comprehension, it is important to rule out the possibility that children constructed perceptual simulations solely as a function of the given task. There are at least three ways in which this might be possible, but none of these possibilities appears to hold.

First, the experiments might have provided the opportunity for drawing backward inferences. That is, the shape of the described object might have been inferred only after viewing the picture probe. This interpretation of the data, however, runs counter to what the mismatch appears to be, namely an effect on picture recognition (see also Hirschfeld & Zwitserlood, 2010) and not on the comparison between the name of the picture and the representation of the sentence. In fact, it is difficult to conceive of a locus of the mismatch effect beyond lexical access to the name of the picture. After all, the name of the picture on which the comparison is based was the same in both conditions.

Second, the effects might be attributed to participants purposefully constructing mental images. Evidence from two previous studies speaks against this possibility. Pecher, van Dantzig, Zwaan, and Zeelenberg (2009) showed that matching sentences facilitated picture recognition more than mismatching sentences when the recognition task was administered 45 min later, unexpectedly, following an unrelated filler task. In a similar vein, Wassenburg and Zwaan (2010) found longer reading times for mismatching sentences, 20 min after participants had viewed a set of pictures, being fully unaware of their relevance to a later reading task. Although the findings from these studies cannot be held as conclusive for the child population in our sample, they provide compelling evidence that language comprehenders retain the shape and orientation of objects.

Third, it might be the case that the involvement of pictures modified the processes operating during or after the sentence presentation. Louwerse and Jeuniaux (2010) showed that the extent to which perceptual representations govern an individual's response on a given task depends on both the instructions and the stimuli that are used. Participants in their study saw pairs of words or pictures that were presented in an iconic vertical configuration (e.g., *attic* above *basement*) or a reverse-iconic vertical configuration (e.g., *basement* above *attic*) and were asked to judge their semantic relatedness or their iconicity. When making semantic relatedness judgments and seeing words as stimuli, response times and error rates were explained best by the order in which the words most frequently occur in language use. When making iconicity judgments or seeing pictures as stimuli, response times and error rates were explained best by the iconicity of the word pair's configuration. One might argue that the use of pictures in the current experiments favored perceptual representations over linguistic representations. However, participants were not asked whether the depicted object provided a good fit to the sentence they read before; they were simply asked whether this object had been mentioned in the sentence. Thus, the use of perceptual representations does not seem to be strongly encouraged in the current experiments. To put our conclusions on more solid footing, future research could investigate whether similar effects can be obtained with verbal-only materials that, in addition, do not involve a judgment task. Nonetheless, there is good reason to believe that the construction of perceptual simulations was spontaneous and did not reflect task-specific processing strategies.

We are now in a position to further discuss the relevance of our findings for developmental theory. An important finding was the absence of an interaction between condition and grade. Even the youngest children showed a mismatch effect, and its size did not increase for children in higher grades. This suggests that the tendency to form perceptual simulations for comprehending sentences such as "Bob saw the pigeon in the nest" and "Bob saw the pigeon in the sky" in perceptual experience is present by the time children enter Grade 2. At the same time, response latencies sharply decreased after Grade 2, indicating that older children are more efficient at performing the task. The lack of an interaction between condition and grade in conjunction with a main effect of condition may be difficult to reconcile with the view that meaning is represented by abstract symbols that are enriched by embodied representations when these become more easily available through experience; rather, children construct simulations of events even if the knowledge of the objects involved is presumably limited.

Another crucial factor constraining the use of perceptual simulations, according to our discussion of the literature, was hypothesized to be processing efficiency (cf. Madden & Zwaan, 2006). This would be reflected in different developmental trajectories for listening and reading given that children's processing resources would be compromised by the reading task. Only with more fluent word reading should the effects in the reading experiment align with those in the listening experiment. An implication of this would be that for children of the same age, listening leads to more perceptual-like situation models than reading. However, the data clearly suggest a different state of affairs. It appears that perceptual simulations are used even when the efficiency of the linguistic processes giving rise to them is still developing. Consistent with this, children as young as 4 years have been found to mentally represent the spatial perspective of characters in narratives (Rall & Harris, 2000; Ziegler, Mitchell, & Currie, 2005) as well as their movement (Fecica & O'Neill, 2010).

Importantly, the current results should not be taken as evidence that the construction of perceptual simulations is equally efficient across all grades studied. In both experiments, participants had 1000 ms to construct a perceptual simulation on the basis of the linguistic input before viewing the picture. Future research could investigate whether shorter intervals are informative as to the time course of the activation and integration of perceptual representations as a function of age.

Overall, the results suggest that perceptual simulations are constructed even when expertise and processing capacity are relatively limited. Perceptual simulation may play a more important role in developing language comprehension than was previously thought. This is the first study to directly address this issue (but see Glenberg et al., 2004, for a comparable discussion). Although no detailed framework exists as of yet in the literature, certain accounts may accommodate these findings. For instance, our interpretation of the results is broadly consistent with theories that place the manipulation of spatial representations at the core of developing cognition (e.g., Mandler, 2010). Under such a view, it is plausible that thought involves the manipulation of perceptual symbols rather than abstract symbols and that language is one of the mechanisms that drives this manipulation. Learning to comprehend language, then, means learning how to use language as a tool for evoking appropriate simulations.

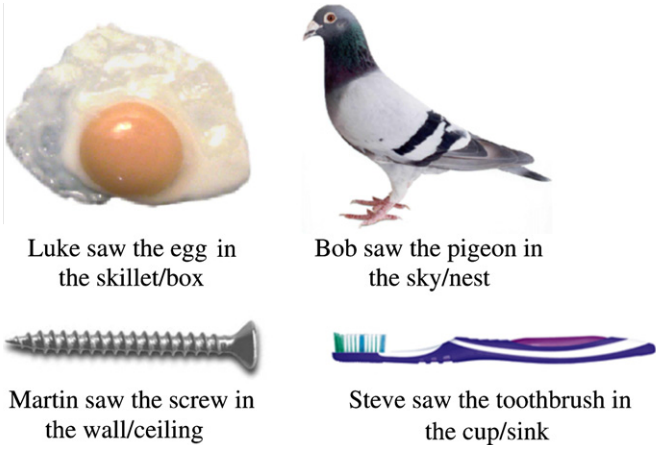
Finally, although our results are supportive of theories of embodied language comprehension, several recent articles have outlined fundamental challenges for such theories (Mahon & Caramazza, 2008; Zwaan, 2009). In particular, the field is in need of research that shows whether embodied representations are essential to comprehension or are epiphenomenal to other processes. While underscoring the importance of addressing these challenges, we believe that the current work holds value in that it extends the descriptive and explanatory power of the simulation view of language comprehension to development during childhood, for both spoken and written materials, and for skilled and less skilled word readers alike.

Acknowledgement

This research was supported by the Board of Public Education Rotterdam (Stichting BOOR). The authors would like to thank the schools and the children for making this research possible.

Appendix A

Samples of experimental sentence–picture pairs



Appendix B

List of experimental sentences (in Dutch, left) and their English translations (right)

Jeroen zag de banaan op het toetje	Jerome saw the banana on the dessert
Jeroen zag de banaan aan de tros	Jerome saw the banana on the bunch
Laura zag de ananas op de taart	Laura saw the pineapple on the pie
Laura zag de ananas in de tas	Laura saw the pineapple in the bag
Jan zag de appel in de fruitsla	John saw the apple in the fruit salad
Jan zag de appel in de rugzak	John saw the apple in the backpack
Tom zag de appeltaart op zijn bordje	Tom saw the apple pie on his plate
Tom zag de appeltaart in de oven	Tom saw the apple pie in the oven
Moeder zag de kip in de oven	Mother saw the chicken in the oven

(*continued on next page*)

Appendix B. (*continued*)

Moeder zag de kip in het hok	Mother saw the chicken in the coop
Jens zag de hond in de mand	James saw the dog in the basket
Jens zag de hond op het grasveld	James saw the dog on the lawn
Vera zag de kat op het hek	Vera saw the cat on the fence
Vera zag de kat in de mand	Vera saw the cat in the basket
Kevin zag het brood in de zak	Kevin saw the bread in the bag
Kevin zag het brood op zijn bord	Kevin saw the bread on his plate
Simon zag het vlees op zijn bord	Simon saw the meat on his plate
Simon zag het vlees in de koelkast	Simon saw the meat in the fridge
Rob zag de duif in het nest	Bob saw the pigeon in the nest
Rob zag de duif in de lucht	Bob saw the pigeon in the sky
Fleur zag de duiker in het water	Flora saw the diver in the water
Fleur zag de duiker op de kant	Flora saw the diver on the beach
Karel zag de eend in de lucht	Carl saw the duck in the sky
Karel zag de eend in de vijver	Carl saw the duck in the pond
Luuk zag het ei in de doos	Luke saw the egg in the box
Luuk zag het ei in de pan	Luke saw the egg in the skillet
Bert zag de spaghetti in de pan	Bert saw the spaghetti in the pot
Bert zag de spaghetti in de verpakking	Bert saw the spaghetti in the wrapping
Ruud zag de danseres in de kleedkamer	Ray saw the dancer in the dressing room
Ruud zag de danseres op het podium	Ray saw the dancer on the stage
Kim zag de handdoek op de plank	Kim saw the towel on the shelf
Kim zag de handdoek aan het haakje	Kim saw the towel on the hook
Liza zag de kaas in de muizenval	Lisa saw the cheese in the mouse trap
Liza zag de kaas op het broodje	Lisa saw the cheese on the bun
Alex zag de ballon in het zakje	Alex saw the balloon in the wrapper
Alex zag de ballon aan het touwtje	Alex saw the balloon on the string
Martijn zag de bloes op de hanger	Martin saw the shirt on the hanger
Martijn zag de bloes op de stapel	Martin saw the shirt on the pile
Kees zag de meloen in de tuin	Keith saw the melon in the garden
Kees zag de meloen in de kom	Keith saw the melon in the bowl
Dirk zag de parasol in de kelder	Dick saw the parasol in the cellar
Dirk zag de parasol op het terras	Dick saw the parasol on the terrace
Vader zag de sigaret in het pakje	Father saw the cigarette in the pack
Vader zag de sigaret in de asbak	Father saw the cigarette in the ashtray
Bas zag de citroen in het drankje	Barry saw the lemon in the drink
Bas zag de citroen op de fruitschaal	Barry saw the lemon in the fruit bowl
Noa zag de tennisster op de stoel	Noah saw the tennis player on the chair
Noa zag de tennisster op de baan	Noah saw the tennis player on the court
Eva zag de tomaat aan de tros	Eve saw the tomato on the bunch
Eva zag de tomaat op de pizza	Eve saw the tomato on the pizza
Lieke zag de ui in de tas	Lea saw the onion in the bag
Lieke zag de ui op de hamburger	Lea saw the onion on the hamburger
Fred zag de auto op de racebaan	Fred saw the car on the race track
Fred zag de auto in de garage	Fred saw the car in the garage
Iris zag de voetballer op het bankje	Iris saw the football player on the bench
Iris zag de voetballer op het veld	Iris saw the football player on the pitch
Stef zag de tandenborstel in de beker	Steve saw the toothbrush in the cup
Stef zag de tandenborstel op de wastafel	Steve saw the toothbrush on the sink
Ruben zag de tak op de stoep	Ruben saw the branch on the curb

(*continued on next page*)

Ruben zag de tak in de grond	Ruben saw the branch in the ground
Maarten zag de schroef in het plafond	Martin saw the screw in the ceiling
Maarten zag de schroef in de muur	Martin saw the screw in the wall
Wouter zag de lamp aan de muur	Walt saw the lamp on the wall
Wouter zag de lamp aan het plafond	Walt saw the lamp on the ceiling
Bart zag de spijker in de vloer	Bart saw the nail in the floor
Bart zag de spijker in het prikbord	Bart saw the nail in the bulletin board
Leo zag de dartpijl in de roos	Leo saw the dart in the bullseye
Leo zag de dartpijl in de vloer	Leo saw the dart in the floor
Moeder zag de vork in de la	Mother saw the fork in the drawer
Moeder zag de vork in de biefstuk	Mother saw the fork in the steak
Roos zag de lepel in de soepkom	Rose saw the spoon in the soup bowl
Roos zag de lepel op het tafelkleed	Rose saw the spoon on the tablecloth
Hugo zag de veer in de inktpot	Hugo saw the quill in the ink jar
Hugo zag de veer in het kippenhok	Hugo saw the quill in the chicken pen
Sara zag de sleutel in het slot	Sara saw the key in the lock
Sara zag de sleutel aan het haakje	Sara saw the key on the hook
Daan zag het boek op de leestafel	Danny saw the book on the reading table
Daan zag het boek in de kast	Danny saw the book in the cabinet
Emma zag de rits aan de tas	Emma saw the zipper on the bag
Emma zag de rits aan het vest	Emma saw the zipper on the waistcoat
Vader zag de pet aan de kapstok	Father saw the cap on the peg
Vader zag de pet op het hoofd	Father saw the cap on the head
Gijs zag de boom op de heuvel	Gus saw the tree on the hill
Gijs zag de boom op de vrachtwagen	Gus saw the tree on the truck

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–609.

Barsalou, L. W., Santos, A., Simmons, W., & Wilson, C. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 245–283). Oxford, UK: Oxford University Press.

Bergen, B., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science, 31*, 733–764.

Brunyé, T. T., Ditman, T., Mahoney, C. R., Walters, E. K., & Taylor, H. A. (2010). You heard it here first: Readers mentally simulate described sounds. *Acta Psychologica, 135*, 209–215.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*, 386–404.

Chiappe, P., Hasher, L., & Siegel, L. S. (2000). Working memory, inhibitory control, and reading disability. *Memory & Cognition, 28*, 8–17.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Dijkstra, K., Yaxley, R. H., Madden, C. J., & Zwaan, R. A. (2004). The role of age and perceptual symbols in language comprehension. *Psychology and Aging, 19*, 352–356.

Fecica, A. M., & O'Neill, D. K. (2010). A step at a time: Preliterate children's simulation of narrative movement during story comprehension. *Cognition, 116*, 368–381.

Fincher-Kiefer, R. (2001). Perceptual components of situation models. *Memory & Cognition, 29*, 336–343.

Gernsbacher, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology, 17*, 324–363.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology, 96*, 424–436.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*, 558–565.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory & Language, 43*, 379–401.

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena, 42*, 335–346.

Hirschfeld, G., & Zwitserlood, P. (2010). How vision is shaped by language comprehension: Top-down feedback based on low-spatial frequencies. *Brain Research, 1377*, 78–83.

Holmes, B. C. (1985). The effect of four different modes of reading on comprehension. *Reading Research Quarterly, 20*, 575–585.

Holt, L. E., & Beilock, S. L. (2006). Expertise and its embodiment: Examining the impact of sensorimotor skill expertise on the representation of action-related text. *Psychonomic Bulletin & Review, 13*, 694–701.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 98*, 122–149.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., et al (2005). Perception of motion affects language processing. *Cognition, 94*, B79–B89.

Kaup, B., Yaxley, R. H., Madden, C. J., Zwaan, R. A., & Lüdtke, J. (2007). Experiential simulations of negated text information. *Quarterly Journal of Experimental Psychology, 60*, 976–990.

Lahey, M. (1988). *Language disorders and language development*. Needham, MA: Macmillan.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*, 259–284.

Long, D. L., Seely, M. R., & Oppy, B. J. (1999). The strategic nature of less skilled readers' suppression problems. *Discourse Processes, 27*, 281–302.

Louwerse, M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition, 114*, 96–104.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109*, 35–54.

Madden, C. J., & Dijkstra, K. (2010). Contextual constraints in situation model construction: An investigation of age and reading span. *Aging, Neuropsychology, & Cognition, 17*, 19–34.

Madden, C. J., & Zwaan, R. A. (2006). Perceptual representation as a mechanism of lexical ambiguity resolution: An investigation of span and processing time. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 1291–1303.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology, 102*, 59–70.

Mandler, J. M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review, 99*, 587–604.

Mandler, J. M. (2010). The spatial foundations of the conceptual system. *Language and Cognition, 2*, 21–44.

Masur, E. F. (1997). Maternal labeling of novel and familiar objects: Implications for children's development of lexical constraints. *Journal of Child Language, 24*, 427–439.

Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory & Cognition, 32*, 1389–1400.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40*, 665–681.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*, 554–566.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, UK: Oxford University Press.

Pecher, D., van Dantzig, S., Zwaan, R. A., & Zeelenberg, R. (2009). Language comprehenders retain implied shape and orientation of objects. *Quarterly Journal of Experimental Psychology, 62*, 1108–1114.

Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.

Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. S. Gorfien (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). Washington, DC: American Psychological Association.

Rall, J., & Harris, P. L. (2000). In Cinderella's slippers? Story comprehension from the protagonist's point of view. *Developmental Psychology, 36*, 202–208.

Richardson, D. C., Spivey, M. J., McRae, K., & Barsalou, L. W. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science, 27*, 767–780.

Rosnow, R. L., & Rosenthal, R. (2005). *Beginning behavioural research: A conceptual primer* (5th ed.). Englewood Cliffs, NJ: Pearson/Prentice Hall.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences, 3*, 417–424.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E–Prime 1.0*. Pittsburgh, PA: Psychological Software Tools.

Shankweiler, D. (1989). How problems of comprehension are related to difficulties in word reading. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 35–68). Ann Arbor: University of Michigan Press.

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

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352–373.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies in discourse comprehension*. New York: Academic Press.

Van Petten, C., Weckerly, J., McIsaac, H. K., & Kutas, M. (1997). Working memory capacity dissociates lexical and sentential context effects. *Psychological Science, 8*, 238–242.

Verhoeven, L. (1995). *Drie-Minuten-Toets [Three-Minute Test]*. Arnhem, The Netherlands: Cito.

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

Yaxley, R. H., & Zwaan, R. A. (2007). Simulating visibility during language comprehension. *Cognition, 105*, 229–236.

Ziegler, F., Mitchell, P., & Currie, G. (2005). How does narrative cue children's perspective taking? *Developmental Psychology, 41*, 115–123.

Zwaan, R. A. (2008). Experiential traces and mental simulations in language comprehension. In M. DeVega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 165–180). Oxford, UK: Oxford University Press.

Zwaan, R. A. (2009). Mental simulation in language comprehension and social cognition. *European Journal of Social Psychology, 7*, 1142–1150.

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 224–245). New York: Cambridge University Press.

Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science, 28*, 611–619.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162–185.
Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*, 168–171.
Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135*, 1–11.