Storybook apps as a tool for early literacy development

Smeets, D.J.H.

Chapter 3

Interactive electronic storybooks for kindergartners to promote vocabulary growth

Abstract

The goals of this study were to examine (a) whether extratextual vocabulary instructions embedded in electronic storybooks facilitated word learning over reading alone and (b) whether instructional formats that required children to invest more effort were more effective than formats that required less effort. A computer-based 'assistant' was added to electronic storybooks. The assistant posed extratextual vocabulary questions. Questions were presented in a multiple-choice format so that children could respond by clicking on the picture that best represented the target word. In Experiment 1 ($N = 20$), children read stories with and without questions. Children learned more words when reading with questions than without. Expressive vocabulary was particularly affected by question insertion. In Experiment 2 ($N = 27$), we used two methods for teaching words: one requiring more effort on the part of children (questions) and one requiring less effort ('hotspots' that provide definitions). Results revealed that questions were more beneficial than just providing a definition or synonym of the target word. Implications for designing new e-book apps are discussed.

Published as:

Smeets, D. J. H., & Bus, A. G. (2012). Interactive electronic storybooks for kindergartners to promote vocabulary growth. *Journal of Experimental Child Psychology, 112,* 36–55. doi:10.1016/j.jecp.2011.12.003

Introduction

It has been repeatedly reported that positive effects of shared book reading on children's vocabulary growth ensue when adults do not just read a story but also pause and pose extratextual questions (e.g., "Where do you see that Bear looks broken-hearted?"). That interactive book reading may be a key to unlocking the full benefits of book sharing for vocabulary growth has been shown in numerous experiments (e.g., Biemiller & Boote, 2006; Blewitt, Rump, Shealy, & Cook, 2009; Elley, 1989; Penno, Wilkinson, & Moore, 2002) and in recent meta-analyses (Mol, Bus, & de Jong, 2009; Mol, Bus, de Jong, & Smeets, 2008). Because young children are increasingly exposed to picture storybooks through DVDs or e-book apps on phones and e-readers, it is a logical step to explore whether extratextual questioning benefits on-screen storybook reading as well (Roskos, Brueck, & Widman, 2009). The experiments presented in the current study focused on factors that may influence the extent to which children benefit and learn from electronically presented books.

Children can learn to associate unknown words with their visual referents via textual exposure to the words in picture storybooks (e.g., Bus, van IJzendoorn, & Pellegrini, 1995). However, children learn 10 to 18% more words when readings include extratextual questions (e.g., Biemiller & Boote, 2006; Blewitt et al., 2009; Brabham & Lynch-Brown, 2002; Collins, 2010). The finding that asking open-ended questions during the storybook reading is effective in promoting expressive word learning (i.e., learning to use a word in an appropriate context) has led to various theories about how to promote expressive word knowledge best (e.g., Ewers & Brownson, 1999; Sénéchal, Thomas, & Monker,1995; Whitehurst et al., 1988). Some believe that retrieval practice (i.e., children verbally reproducing words) is crucial (e.g., Sénéchal, 1997), whereas others assume that questions encourage children to be more engaged in extracting meaning (e.g., Ewers & Brownson, 1999).

Analogous to traditional shared book reading, electronic storybook exposure has been demonstrated to support kindergartners' vocabulary (e.g., Korat, 2010; Shamir, Korat, & Barbi, 2007; Smeets & Bus, 2012a; Verhallen & Bus, 2010; Verhallen, Bus, & de Jong, 2006; see also Zucker, Moody, & McKenna, 2009 , for a meta-analysis of studies published before 2009). In addition, there is evidence that interactive e-books with built-in vocabulary instructions promote word learning, yet these studies are limited in their design. First, a read-only control

condition is often not included (e.g., Higgins & Cocks, 1999; Segers, Takke, & Verhoeven, 2004; Segers & Verhoeven, 2002, 2003; Shamir, Korat, & Shlafer, 2011). Thus, the studies do not isolate the additional effects of extratextual instructions over e-book reading alone. Second, because the studies involved experimenters sitting next to and interacting with children, positive effects might have been the result of the presence/monitoring of adults rather than the programs themselves (Higgins & Cocks, 1999; Higgins & Hess, 1999).

In this article, we report on two experiments in which we examined whether extratextual vocabulary instructions are a valuable addition to electronic storybook reading. The goal of Experiment 1 was to test whether children learn more words from electronic storybook readings when stories include multiple-choice (MC) vocabulary questions than when they do not include questions. In this experiment, difficult words that were encountered in the story text were highlighted by questions modeled on adult–child interactions. Questions were designed in an MC format followed by individualized feedback. For instance, when 'Bear' (a character in the story Beer is op Vlinder [Bear is in love with Butterfly]; van Haeringen, 2004) 'fans' a fire, the story is interrupted and the child is asked a question about 'fanning' (e.g., "In which picture does Bear fan the fire?"). Children respond to the question by selecting the correct image of 'Bear fanning the fire' among two distracters (e.g., 'Bear sitting next to the fire' and 'Bear destroying a house'). Modeled on adult–child extratextual conversations, children received feedback on their answers. The feedback often included synonyms or definitions (e.g., "Good job! Here you see Bear fanning the fire; the fire is growing"). Incorrect responses were first followed by incentives to try again (i.e., the question was repeated). After a second incorrect response, cues were provided to help the child find the correct solution (e.g., "How does Bear make sure that the fire keeps on burning?"). After a third incorrect response, the 'computer pal' modeled the correct answer and explained why this image represented the target word best. To examine the additive value of MC questions, we tested whether target words were learned more often when children received questions than when not.

In Experiment 2, we tested whether the effectiveness of vocabulary instruction was enhanced by active participation on the part of children. We contrasted the questions in Experiment 1 with an alternative way of instructing vocabulary as may typically occur in interactive e-books: some scenes of the story include 'hotspots' that provide a definition of the object when clicked on, a technique

similar to strategies that are used during whole group book reading in classrooms (Biemiller & Boote, 2006). For instance, when children click on Grandpa who is 'taking a nap' (Boonen, 2004), they hear a definition of Grandpa's target behavior: "Grandpa is taking a nap; he is sleeping for a little while." This experiment tests whether the computer-assisted question format (our adult substitute) was more effective than only hearing a definition of a target word in context. In other words, does selecting the best-fitting picture among two distracters provide additional benefits over simply hearing the definition of the word?

In the current experiments, knowledge of target words was tested at different levels (i.e., receptive and expressive) because previous research has suggested that word knowledge is acquired in degrees (e.g., De Temple & Snow, 2003; Ouellette, 2006). Nagy and Scott's (2000) principle of incrementality suggests that word knowledge exists on a continuum from no knowledge to varying levels of partial knowledge to a more complete understanding of a word's meaning. When children first become familiar with a word, they may only be able to select its depiction among alternatives, but when the understanding of the word is fine-tuned, they may actively use that word in an appropriate context. In this perspective, it seems plausible to assume that receptive vocabulary knowledge (identifying a word's visual referent among alternatives) may be the forerunner of expressive vocabulary knowledge (retrieving a word from memory and using it in an appropriate context) (Chan, Cheung, Sze, Leung, & Cheung, 2008; Laufer & Paribakht, 1998; Stahl & Stahl, 2004; Verhallen & Bus, 2010). Examining word learning in the perspective of prior knowledge as we did in the current experiments may enlarge our understanding of vocabulary development.

More than textual exposure to words alone, questions may promote associations of unknown words with their visual referents and lead to growth in receptive vocabulary knowledge. However, questions might not automatically result in more expressive knowledge; that is, children may understand 'chaos,' but the word might not be sufficiently familiar to children to use the word in an appropriate context and pronounce it correctly. Studies on adult–child book reading suggest that expressive word knowledge can be stimulated best with 'two-way' interactive reading sessions in which parents pose questions during the storybook reading (e.g., Ewers & Brownson, 1999; Sénéchal et al., 1995; Whitehurst et al., 1988). If outcomes found in these experiments depend on children verbally reproducing words (Sénéchal, 1997), the MC questions in the current experiments might not promote expressive knowledge to the

same extent. Children respond by mouse clicking rather than verbally, which may keep children from fine-tuning their understanding of a word and storing information about the word's pronunciation (Sénéchal, 1997). If, however, questions during adult–child interactions are effective because they encourage children to be more engaged in extracting meaning (e.g., Ewers & Brownson, 1999), effects of MC questions on expressive vocabulary growth may be similar to those of adult questioning. Given that repeatedly listening to verbal stimuli has been demonstrated to activate speech-related motor centers in the brain (Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Liberman & Mattingly, 1985; Rizzolatti & Craighero, 2004; Watkins, Strafella, & Paus, 2003), practice of the phonological structure of words resulting in expressive word learning may still occur.

In sum, the experiments reported here tested (a) whether children learned more words from e-books when the books included questions about words and (b) whether active child participation enhanced vocabulary learning from e-books.

EXPERIMENT 1

Prior experiments have tested effects of interactive electronic storybooks on either receptive or expressive knowledge (Higgins & Cocks, 1999; Higgins & Hess, 1999; Segers & Verhoeven, 2003). This experiment tested effects of extratextual vocabulary instructions on both levels of word knowledge. Effectiveness of MC questions modeled on adult–child interactions was tested by contrasting exposure to target words in the text alone with words that were instructed via MC questions. A second aim of this experiment was to test potential disadvantages of an interruptive reading style (e.g., Blewitt et al., 2009; Dickinson & Smith, 1994; Reese & Cox, 1999) by contrasting effects of questions interspersed throughout the story text with questions posed at the end of the session (after the entire story has been read without interruptions). For electronic storybooks in particular, interruptions have been reported to disrupt children's learning behavior, perhaps because interactive features interrupt the flow of the story (e.g., de Jong & Bus, 2002; Labbo & Kuhn, 2000; Trushell & Maitland, 2005). In contrast to interactive moments in the previous studies, however, in the current research the interruptions in the e-book apps were of short duration and the story

continued immediately after the question was answered. However, it is possible that vocabulary instruction is more beneficial when questions are placed at the end of the story, thereby not interrupting the flow of reading.

First, we may expect that answering questions at the end of the story requires more effort than answering questions interspersed throughout the story text, resulting in more vocabulary growth (Cennamo, 1993). Questions that are posed immediately after the target word is encountered in the text may be easier to answer because it may suffice to simply click on the picture that was visible in the preceding scene without in-depth processing of the alternative answers. This might result in more errors in questions afterward than in questions throughout the story. Second, questions interspersed throughout e-books may easily exceed the limited processing capacity of the cognitive system (Baddeley, 1998). According to the theory of multimedia learning, this may cause cognitive overload and interfere with learning (Baddeley, Gathercole, & Papagno, 1998; Gathercole, Service, Hitch, Adams, & Martin, 1999; Mayer & Moreno, 2003; Reed, 2006). If questions are a threat to learning, we may expect negative effects on uninstructed words; children may acquire more uninstructed words when books do not include questions at all or with questions at the end.

In sum, Experiment 1 tested (a) whether answering MC questions increases receptive and expressive word learning over e-book reading alone and (b) whether the benefit of MC questions varies as a function of whether they are presented throughout the story or only after the story has been read completely.

Method

Participants

A total of 20 junior kindergartners (11 boys and 9 girls) between 4 and 5 years of age (M = 54.50 months, SD = 2.52) participated in the current experiment. Participants were selected from two Dutch primary schools and were typically developing children from middle-socioeconomic status (SES) families with Dutch as their first language. Participating children's mean standardized scores on the Peabody Picture Vocabulary Test (PPVT, M = 108.40, SD = 13.86) confirmed that the sample was average in vocabulary.

Design

Conditions. A pretest–posttest within-subjects design was used to examine the additive value of MC questions and the effect of timing. Each child read five electronic stories; one story was presented in a read-only control condition without MC questions (read-only); two stories were each interrupted four times for MC questions (MCQ during); and two stories were presented without interruption, each with four MC questions at the end of the story (MCQ after). Because previous research has not revealed measurable effects on vocabulary after only one exposure to the story, we decided that children would read all stories twice (e.g., Verhallen et al., 2006).

Each child read stories in every condition. Because all five stories were used as stimulus material in each of the three conditions, we created unique combinations of stories and conditions for each of the 20 participants (see Table 3.1 for an example). Importantly, each of the five stories appeared with equal frequency within each condition so that any effect of condition would not be the outcome of differences in stories.

Table 3.1. Example of how stories were assigned to conditions in Experiment 1.

	Story A	Story B	Story C	Story D	Story E
child 1	Read-Only	*MCQ During*	*MCQ During*	***MCQ After***	***MCQ After***
child 2	***MCQ After***	Read-Only	*MCQ During*	*MCQ During*	***MCQ After***
child 3	***MCQ After***	***MCQ After***	Read-Only	*MCQ During*	*MCQ During*
child 4	*MCQ During*	***MCQ After***	***MCQ After***	Read-Only	*MCQ During*
child 5	*MCQ During*	*MCQ During*	***MCQ After***	***MCQ After***	Read-Only

Target words. Per book, we selected eight target words that were pre- and posttested (see Appendix A). All of these were low-frequency words according to Schrooten and Vermeer's (1994) analyses of various bodies of verbal contexts in kindergarten classrooms (e.g., storybooks, teacher–child talk). Frequencies of selected words ranged between 0 and 15, meaning that the majority of kindergartners would not be familiar with these words.

All target words appeared in the story text and could either be additionally instructed with MC questions or remain uninstructed, meaning that no question was posed about the words. In the read-only story, all target words remained uninstructed. When stories included MC questions (both during and after the

story), half of the target words encountered in that condition were instructed and the other half were not.

To ensure that all target words in the stories were instructed as often as they remained uninstructed in both of the questioning conditions (MCQ during and MCQ after), we created two sets of words per story. Of the 20 participants, 10 received Set 1 as the instruction set in all four stories with questions (meaning that words in Set 2 remained uninstructed), whereas the other 10 participants received Set 2 for instruction (with words in Set 1 remaining uninstructed).

Materials

Intervention materials.

Storybooks. Five Dutch storybooks were available as electronic storybooks: Beer is op Vlinder [Bear is in Love with Butterfly] (van Haeringen, 2004), Rokko Krokodil [Rokko the Crocodile] (de Wijs, 2001), Bolder en de Boot [Bolder and the Boat] (Hoogstad, 2005), Met Opa op de Fiets [Cycling with Grandpa] (Boonen, 2004), and Tim op de Tegels [Pete on the Pavement] (Veldkamp, 2004). Each of the stories was available in all three conditions (read-only, MCQ during, and MCQ after).

In the current experiment, we used video versions of these five picture storybooks. The static illustrations in the original books were transformed into motion pictures that detailed story events in the oral text. For instance, in Bear is in love with Butterfly (van Haeringen, 2004), Bear tries to confess his love to Butterfly but fails hopelessly because he is very shy. In the video storybook, not only do we see how Bear looks when he is shy (red cheeks and not daring to look up), but also the animated illustration shows the transformation of how Bear's cheeks turn red while he talks to Butterfly and how he bows his head more and more. Simultaneously, we hear Bear mumbling and music in the background. The oral text is exactly the same as the text in the print version of the books. However, the print was not included in the storybooks we used in the current experiment.

Computer pal. In all story formats (with or without questions), a computer pal introduced the storybook: "Hi! Nice to see you! We are going to read a story together." In questioning conditions, the computer pal added, "Sometimes, I will ask you a question; you have to use the computer mouse to answer it." When questions interrupted the reading sessions, the computer pal would pop up after a scene of the story and say, "Time for a question!" During questions,

the computer pal gave appropriate feedback to the child's response. After the question was answered, the computer pal announced: "Now, we will continue with the story." Questions were introduced similarly when they were inserted after the storybook reading. At the end of the session (regardless of the story format), the computer pal would conclude by saying, "That was fun! Next time, we will read another story. See you then!"

Questions. A computer pal posed four MC vocabulary questions either during or after the storybook reading. In the MCQ during condition, the questions appeared directly after relevant passages in the story and the story continued after the question had been answered correctly. In the MCQ after condition, children watched the story without interruptions and all four questions were posed after the entire story had been read.

Multiple-choice format. Three pictures representing alternative answers appeared on-screen after the computer pal had posed the question (see Fig. 3.1 A). For instance, in Bear Is in Love with Butter y (van Haeringen, 2004), Bear is shy. The computer pal asks, "Bear is shy, where can you see that?" after which three pictures (originating from the story) are presented on-screen: (a) Bear being broken-hearted, (b) Bear being shy, and (c) Bear being angry. All three alternatives originate from illustrations or details of illustrations in the same storybook. Children answered the question by clicking on one of the three pictures.

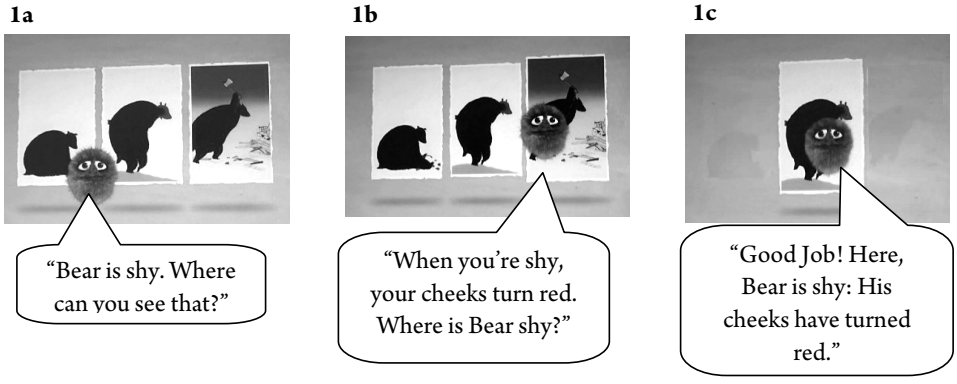

Figure 3.1. Screenshots of a multiple choice question. *Figure 1a*: the computer pal asks a question and shows three alternative answers on screen. *Figure 1b:* after an incorrect response the computer assistant provides feedback including a clue. *Figure 1c*: after a correct response the assistant explains why the answer is correct. The computer pal's feedback, here presented in the callouts, is actually provided orally.

Feedback. Errors in answering questions were followed by increasingly supportive feedback. First the question was repeated, next a clue was given (e.g., "When you're shy, your cheeks turn red. Where is Bear shy?"; see Fig. 3.1 B), and finally the computer pal demonstrated the correct picture while attracting attention to relevant details (e.g., "Here, Bear is shy, his cheeks have turned red."). Correct responses were followed by confirmation of the correct responses (see Fig. 3.1 C). Because stories were presented through a Web portal, all feedback could be registered in an online data store. We could derive the number of errors children made per question from the amount of feedback that was provided. For instance, one feedback clue meant that children made one error in answering that particular question, which also led to one extra repetition of the target word.

Tests.

Peabody Picture Vocabulary Test. For screening purposes only, children's general receptive vocabulary was assessed using the PPVT-III-NL (third edition, Dutch; Schlichting, 2005). The main aim was to verify that children's language scored in the normal range.

Knowledge of target words. All 40 target words were pre- and posttested receptively and expressively:

(a) *Receptive target vocabulary.* To assess children's receptive knowledge of word meanings, children were asked to select the target word out of four pictures. The correct image was presented among three distracters, all of which were selected from the same storybook. The test consisted of 40 items (all 40 target words).

(b) *Expressive target vocabulary*. A cued expressive vocabulary task in which children were asked to complete sentences with target words was designed. While the experimenter read an incomplete stimulus sentence aloud (e.g., "Here you can see that Bear is..."), corresponding pictures from the storybooks (e.g., Bear being 'broken-hearted') were shown on-screen. All 40 words were tested in this way with cue sentences that did not resemble the exact phrases in the target stories. Children's responses were coded as correct when they completed the sentence with the target word. When children used a synonym (e.g., 'sad' instead of 'broken-hearted'), the response was coded as incorrect. Intraclass correlation equaled .996.

Procedure

Testing and intervention took place in a separate room at children's school. Preceding the intervention, three sessions lasting approximately 10 min were preserved for administering the PPVT (as a screening tool) and target vocabulary tests. Because the receptive vocabulary test could influence scores on the expressive vocabulary test; the latter was always administered first.

Intervention was carried out during five sessions over the subsequent 2.5 weeks, during which each of the five stories was read twice. There were two sessions per week, with two stories being presented each session. A within-subjects design was applied, with each child being exposed to all three conditions (read-only, MCQ during, and MCQ after). Over sessions, the target stories moved up. For instance, in the first session Child 1 read Stories A and B, in the second session Stories C and D, in the third session Stories E and A, and so on. Child 2 started with Story B and moved up over sessions in the same way as Child 1. Thus, four other stories were presented before a child encountered the same story again, and the order of stories varied between participants. Because there were no rules concerning the format of two stories presented during one session, children might read stories assigned to the same or different conditions during a session (see Table 3.1).

Two children worked simultaneously in the computer room. Children used headphones to prevent them from disturbing each other and to facilitate working independently. After children were logged into the Web portal by the experimenter, the experimenter did not interfere. Only in the rare instances when children were very distracted (e.g., playing with a toy in their pocket, continuously trying to engage the experimenter in conversation), the experimenter encouraged them to attend to the story. All children were able to use the computer mouse to answer the MC questions.

Next, 2 to 3 days after the last intervention session, children's expressive and receptive target vocabulary was posttested in two separate sessions. The receptive task was administered after the expressive vocabulary test.

Results

An overview of pre- and posttest scores in each condition is provided in Table 3.2. To test the difference between pretest and posttest, we performed nonparametric Wilcoxon tests because normality assumptions were not satisfactory for pretest scores for both receptive and expressive vocabulary. To correct the familywise Type I error, we tested at a significance level of .01 ($\alpha = .05/5$ comparisons). For expressive vocabulary, increases were significant in every condition (all $ps < .001$). For receptive word learning, 4 of 5 increases were significant ($ps < .01$). In further analyses, we primarily used gain scores (i.e., the difference between pretest and posttest scores) as a dependent variable because these scores were distributed normally.

Table 3.2. Means (and sd's) for instructed and uninstructed vocabulary (receptive and expressive) in three conditions in Experiment 1.

		Receptive		Expressive	
		Instructed	Uninstructed	Instructed	Uninstructed
Read-only	Pre		3.80 (1.44)		0.30 (0.57)
	Post		4.70 (1.89)		1.95 (1.39)
MCQ during	Pre	3.55 (1.54)	3.90 (1.48)	0.55 (0.83)	0.55 (0.69)
	Post	6.45 (1.61)	5.20 (1.28)	3.35 (1.35)	1.70 (1.17)
MCQ after	Pre	3.80 (1.44)	3.95 (1.39)	0.40 (0.60)	0.55 (0.83)
	Post	6.30 (1.17)	5.25 (1.74)	2.85 (1.42)	1.75 (1.16)

Note. Maximum score equals 8 for all variables.

Overall effects of MC questions

To examine whether interspersing questions throughout a story (MCQ during condition) impaired learning of uninstructed words compared with storybook readings without interruptions (MCQ after and read-only conditions), we conducted two separate repeated-measures analyses of variance (ANOVAs) on receptive and expressive gains in uninstructed words. In both analyses, condition (read-only, MCQ during, or MCQ after) was used as a within-subjects factor. We found no effect for condition on either receptive or expressive vocabulary

(both *p*s > .31), indicating that interrupting the flow of reading did not interfere with learning uninstructed words (see Fig. 3.2).

To examine whether MC questions promoted word learning, we compared vocabulary gains for instructed and uninstructed words in both questioning conditions (MCQ during and MCQ after), expecting superior gains for instructed words. We conducted a repeated-measures ANOVA on gain scores using word type (instructed or uninstructed), condition (MCQ during or MCQ after), and level of word knowledge (receptive or expressive vocabulary) as within-subjects factors. Results demonstrated superior gains for instructed words rather than uninstructed words, F (1, 19) = 37.17, p<.001, η_p^2 =.66, d = 1.77, but effects for condition (MCQ during or MCQ after) and level of word knowledge (receptive or expressive vocabulary) were nonsignificant (all *p*s > .47). All interactions among condition, word knowledge, and level of word knowledge were nonsignificant as well (*p*s > .36). Thus, there was a significant effect of questions on instructed word learning in both questioning conditions (MCQ during and MCQ after) and in both receptive and expressive vocabulary, as displayed in Fig. 3.2.

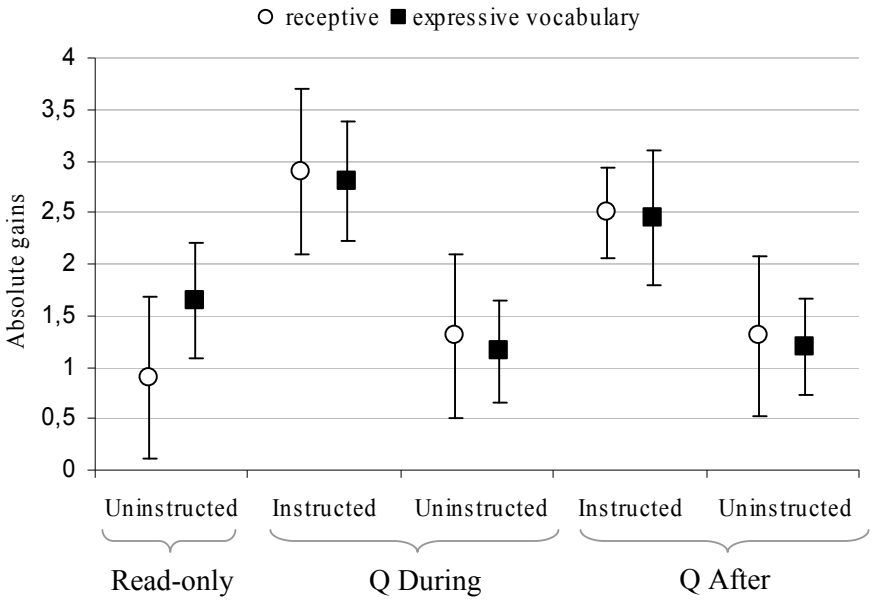

Figure 3.2. Receptive and expressive vocabulary gains (with confidence intervals) in Experment 1, displayed per condition and word test. Gains are calculated by subtracting pretest from posttest scores with maximum scores of 8 per variable.

Effects on novel vocabulary

Next, we examined the effect of questions on words that were completely unfamiliar at both pretests. For each child, we examined each novel word separately and coded whether it had been learned receptively, expressively, both ways, or not at all. Then, we calculated what percentages of completely unknown instructed and uninstructed words were learned receptively, expressively, and both ways. Scores for MCQ during and MCQ after were pooled because there were no differences between these two conditions. Because normality assumptions were not satisfactory, a nonparametric test (Wilcoxon) was used rather than a paired-samples t-test for comparing instructed and uninstructed words.

Half of all target words (52.5% of instructed and 48.1% of uninstructed words) were completely unknown at pretest. Analyses revealed that children learned more instructed than uninstructed words, $z = -3.53$, $p < .001$, $d = 2.60$. Figure 3.3 displays which percentages of novel words were learned receptively only, expressively only, or both ways (receptively and expressively), for instructed

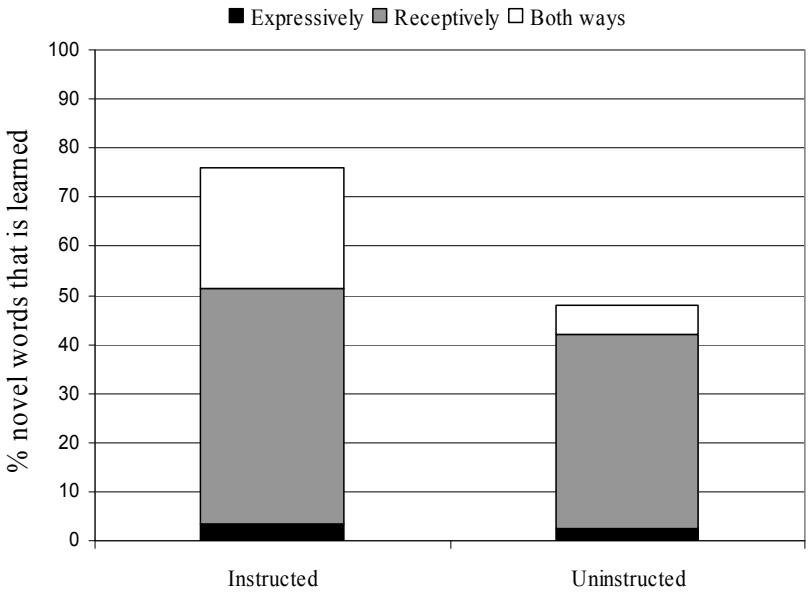

Figure 3.3. Percentage of instructed and uninstructed words unfamiliar at pretests that were learned expressively only (black), receptively only (grey), or both ways (white) in Experiment 1.

and uninstructed words separately. In line with the prior finding that receptive word knowledge precedes expressive word learning (Verhallen & Bus, 2010), a very small percentage of words was learned expressively only. The majority of words were learned receptively only, with higher gains for instructed (48%) than for uninstructed words (40%), $z = -2.25$, $p = .02$, $d = 1.15$. The percentage of words also learned expressively was substantially larger for instructed words (24%) than for uninstructed words (6%), $z = -3.25$, $p < .001$, $d = 2.14$. Thus, 13 percent of uninstructed words that were learned receptively were also learned expressively (6% relative to a total of 46%), whereas 33% of instructed words that were learned receptively were additionally learned expressively (24% relative to a total of 72%).

Effects on partially familiar vocabulary

Next, we selected all words that children knew receptively but not expressively at pretest (41.3% of instructed and 42.8% of uninstructed words) and coded what percentage of these partly familiar words was learned expressively (at posttest). In line with the findings reported above, the percentage of words that was learned expressively at posttest was significantly larger for instructed (45%) than for uninstructed words (20%), $z = -3.46$, $p < .001$, $d = 2.41$. Thus, questions also promoted expressive knowledge for words that were receptively familiar before the intervention.

Errors

The number of times that feedback was provided per question reflects the number of errors. In the MCQ during condition, children made on average .33 errors ($SD = .23$) per question. In the MCQ after condition, children made significantly more errors, $t(19) = -2.35$, $p = .03$, $d = .65$, with an average of .53 ($SD = .37$).

Discussion

MC questions were found to significantly contribute to children's vocabulary gains with impressive effects on instructed words (Cohen's ds > 1.50 for both receptive and expressive vocabulary). Children learned approximately 15% of target words that were encountered in the text with no additional instruction,

with similar overall gains made for receptive and expressive vocabulary. MC questions added another 18% gain to both receptive and expressive vocabulary (amounting to an average gain of 33%), which is comparable to the reported additive value of adult questions during adult–child book sharing (e.g., Biemiller & Boote, 2006; Brabham & Lynch-Brown, 2002; Collins, 2010). The current result that MC questions stimulate expressive vocabulary is incompatible with the hypothesis that pronunciation of target words is essential for expressive word learning (Sénéchal, 1997). This finding is consistent with Walsh and Blewitt's (2006) finding that expressive knowledge improves as much when questions do not require children to use target words as when they do. Interestingly, MC questions promoted expressive word knowledge not only when words were receptively known from the outset but also when words were completely novel to children. Therefore, we may conclude that questions supported both retrieving words in an appropriate context and actively using and pronouncing words correctly.

In line with Blewitt and colleagues' (2009) findings, timing of questions seemed unimportant. Children did not benefit more from questions afterward, although more errors at the end of the text indicated that those questions were more challenging than questions that interrupted the story. Nor did we find support for the hypothesis that interruptions overloaded children's verbal short-term memory, thereby interfering with deriving meaning of unknown words from text (Mayer & Moreno, 2003); our findings show no evidence that children learn more uninstructed words when the flow of reading is not interrupted.

EXPERIMENT 2

A unique finding of our first experiment is that MC questions posed by a computer assistant promoted not only receptive but also expressive knowledge of novel words. A plausible explanation is that questions require children to actively participate by digesting the question, considering the possible answers, and finally deciding which alternative represents the target word best. As Salomon (1984) hypothesized nearly three decades ago, increases in the amount of invested mental effort (AIME) may result in more learning. Studies on adult–child book sharing have suggested that even though expressive word learning can benefit from an adult labeling and/or defining words (e.g., Biemiller

& Boote, 2006; Elley, 1989; Penno et al., 2002), children may learn more words when instruction requires higher levels of involvement and participation (Ewers & Brownson, 1999; Sénéchal, 1997; Sénéchal et al., 1995). Questions may facilitate in-depth processing, thereby promoting deeper meanings of words and fine-grained semantic differentiation (McKeown, Beck, & Apthorp, 2011).

A main purpose of the second experiment was to compare extratextual vocabulary instructions with low and high involvement. As an alternative for MC questions, we designed vocabulary instructions for electronic storybooks similar to Biemiller and Boote's (2006) definitions during whole-group readings. In this alternative instruction, we used the same definitions as provided in the MC questions but provided the information rather than asking a question about the word and expecting children to answer. After a target word was encountered in the text, the animated illustration froze and the mouse changed into a magnifying glass that children could move over the picture in search for a hotspot. After clicking on a hotspot item (recognizable by a green border), children heard a definition of the target object, action, or quality. For instance, the story Pete on the Pavement (Veldkamp, 2004) freezes when Pete walks out of his house 'by himself.' After clicking on Pete in the illustration, the expression 'by himself' is repeated and defined ("Pete is going outside by himself: all alone"). The story continues until the next target word is encountered and the animated illustration freezes again.

In this second experiment, children independently read a series of digital video storybooks either with no vocabulary instructions (read-only), with hotspots that included word meaning explanations (low level of involvement), or with MC questions (high level of involvement). Children may benefit more from MC questions than from hotspots due to higher levels of involvement. Because there is evidence suggesting that increased mental effort may provide a strong link for memorizing novel words in particular (Ewers & Brownson, 1999), the level of children's involvement should have differential effects on novel words (from the outset not known receptively or expressively) versus partially familiar words (from the outset known receptively but not expressively). In-depth processing is expected to be more crucial for learning novel words than for elaborating knowledge of partly familiar words.

Because the feedback procedure for questions entails target words being repeated more often when children make more errors, we wondered whether positive effects of questioning may result from hearing words more often

(Blewitt et al., 2009; Justice, 2002). Therefore, we also examined whether more repetitions of target words in the questioning condition predict word learning.

Method

Participants

Participants were 27 junior kindergartners (13 boys and 14 girls) between 4 and 5 years of age (M = 57.56 months, SD = 3.68). Participants were selected from three Dutch primary public schools and were typically developing children from middle-SES families with Dutch as their first language. According to children's mean standardized scores on the PPVT (M = 105.52, SD = 12.42), the sample's language abilities were average. None of these children had participated in Experiment 1.

Design

Conditions. We used a pretest–posttest within-participant design to examine differential effects of two kinds of interactive vocabulary instructions (MC questions vs. hotspots with word definitions). Each child read five electronic video stories twice, of which (a) one story was presented in a read-only condition without vocabulary instructions (read-only), (b) two stories were interrupted four times for MC questions about difficult words (questions), and (c) two stories were interrupted four times for hotspots with word definitions or synonyms (hotspots).

Each child read stories in every condition. As in Experiment 1, we created a unique combination of stories and conditions for each of the 27 participants. Importantly, each of the five stories appeared with equal frequency within each condition.

Target words. In this experiment, we used the same target words as in Experiment 1 (eight per book). In the read-only condition, all target words were uninstructed (n = 8). For stories in instruction conditions (questions and hotspots), half of the target words were instructed (four per book, eight per condition) and the other half remained uninstructed (four per book, eight per condition). Similar to Experiment 1, we created two word sets. We ensured that each word was instructed as often as it remained uninstructed (see Appendix A).

Materials

Intervention materials.

Storybooks. The same storybooks as in Experiment 1 were used. All stories were available as read-only, with questions, and with hotspots.

Questions. The questions condition in the current experiment was similar to the MCQ during condition in Experiment 1 (see Fig. 3.1). After a computer pal had interrupted the oral rendition of text for an MC question, three pictures representing alternative answers appeared on-screen. A feedback procedure started if the question was not answered correctly. The computer pal always provided a definition or synonym of the target word after the question had been answered.

Hotspots. Four times during a session, a scene of the story froze and the mouse changed into a magnifying glass that children could move over the picture, searching for the target detail (see Fig. 3.4 for an example). When the mouse skimmed the hotspot, the object was marked in green. Clicking on the hotspot resulted in enlargement of the object combined with an oral explanation. For instance, the story Pete on the Pavement (Veldkamp, 2004) froze when Pete walked out of his house, and after clicking on the depiction of the paving stones, the word 'paving stones' was repeated and defined. The story continued until the next target word appeared. During 'hotspot moments', children had a chance to explore the frozen picture for 30 s. In rare cases where children failed to click on the hotspot, the magnifying glass would automatically turn to the hotspot, thereby highlighting the target object and revealing a definition. Target words were defined in exactly the same way as in the questions condition.

Tests. Receptive knowledge (selecting the target word out of four pictures) and expressive knowledge (completing stimulus sentences) of all 40 target words were pre- and posttested similar to Experiment 1.

Procedure

Before the intervention, the PPVT was administered as a screening tool. Target vocabulary tests (expressive first) were administered as pretests. During a 2- to 3-week intervention with two sessions per week, all five stories were read twice. Each child read stories in each condition (read-only, questions, and hotspots). The order of the stories differed for all children (see Experiment 1 for a precise description of the procedure). All children used headphones and worked independently. Children used the computer mouse for answering MC questions or clicking on hotspots. Next, 2 to 3 days after the last intervention session, expressive and receptive target vocabulary was posttested.

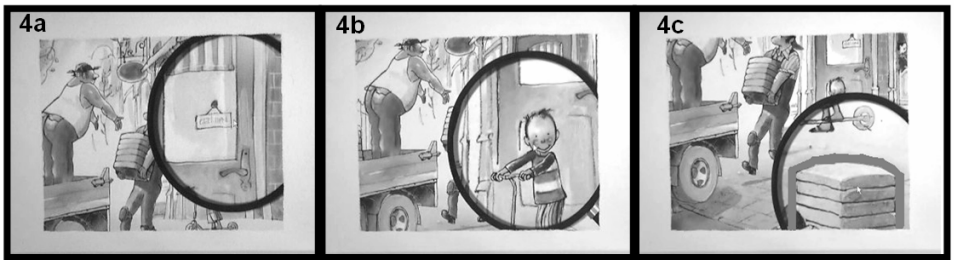

Figure 3.4. During a 'hotspot-moment' a looking glass appears that can be moved over the illustration while it magnifies what is under it; see for instance the tag fixed at the door (4a) and Pete (4b). When the looking glass moves over the hotspot, here *paving stones*, it colors green (4c). Clicking on the green encircled hotspot results in a repetition and definition of the target word (in this case: "Here you see paving stones. Paving stones are used for making the pavement.").

Results

An overview of pre- and posttest scores for all three conditions is presented in Table 3.3. To test the difference between pretest and posttest, we performed several nonparametric Wilcoxon tests in each condition because normality assumptions were not satisfactory for pretest scores for both receptive and expressive vocabulary. To correct the familywise Type I error, we used a .01 significance level ($\alpha = .05/5$ comparisons). All increases were significant (all ps $< .001$).

Table 3.3. Means (and sd's) for instructed and uninstructed vocabulary (receptive and expressive)in three conditions in Experiment 2.

		Receptive		Expressive	
		Instructed	Uninstructed	Instructed	Uninstructed
Read-only	Pre		4.11 (1.40)		0.33 (0.68)
	Post		5.63 (1.60)		1.59 (1.37)
MC questions	Pre	3.78 (1.28)	3.41 (1.47)	0.48 (0.75)	0.37 (0.56)
	Post	5.52 (1.40)	5.30 (1.75)	3.59 (1.55)	1.48 (1.19)
Hotspots	Pre	2.93 (1.49)	3.59 (1.72)	0.44 (0.64)	0.26 (0.45)
	Post	5.26 (1.93)	5.37 (1.67)	2.37 (1.47)	1.48 (1.19)

Note. Maximum score equals 8 for all variables.

Questioning conditions versus read-only

The goal of our first analysis was (a) to replicate Experiment 1's finding that interspersing questions throughout a story does not impair uninstructed word learning and (b) to examine whether definitions (hotspots condition) are equally noninterfering. We performed a repeated-measures ANOVA on gains in uninstructed words with condition (read-only, questions, or hotspots) as a within-subjects factor. Normality assumptions were satisfactory for gain scores. No effects were found for condition on either receptive or expressive vocabulary (both *p*s > .71). This result confirms the finding in Experiment 1 that interrupting the reading to focus attention on target words (with either questions or hotspots) did not interfere with learning other words from the text.

MC questions versus hotspots

A second repeated-measures ANOVA was performed on gain scores using word type (instructed *vs.* uninstructed words), condition (questions *vs.* hotspots), and level of word knowledge (receptive *vs.* expressive vocabulary) as within-subject factors. Children learned significantly more instructed than uninstructed words, $F(1, 26) = 22.16, p < .001, \eta_p^2 = .46, d = 1.04$. There was no main effect of receptive *vs.* expressive vocabulary, $p = .66$, yet the interaction between word type and word knowledge was significant, $F(1, 26) = 6.11, p = .02, \eta_p^2 = .19, d = .39$. The three-way interaction between condition, word type, and word knowledge was also significant, $F(1, 26) = 6.32, p = .02, \eta_p^2 = .20, d = .40$).

In a post hoc analysis on receptive vocabulary, neither main nor interaction effects were found for condition and word type, all *p*'s > .25, indicating that MC questions and hotspots were as effective as reading alone. For expressive vocabulary, on the other hand, a repeated measures ANOVA revealed a significant main effect of word type (instructed *vs.* uninstructed words), $F(1, 26) = 47.68$, $p < .001$, $\eta_p^2 = .65$, $d = 1.70$. Although the main effect for condition (questions *vs.* hotspots) did not reach significance, $F(1, 26) = 3.65$, $p = .07$, $\eta_p^2 = .12$, $d = .25$) there was a significant interaction between condition and word type, $F(1, 26) = 10.57$, $p = .003$, $\eta_p^2 = .29$, $d = .60$. As can be seen in Figure 3.5, MC questions were more effective in promoting instructed words than hotspots.

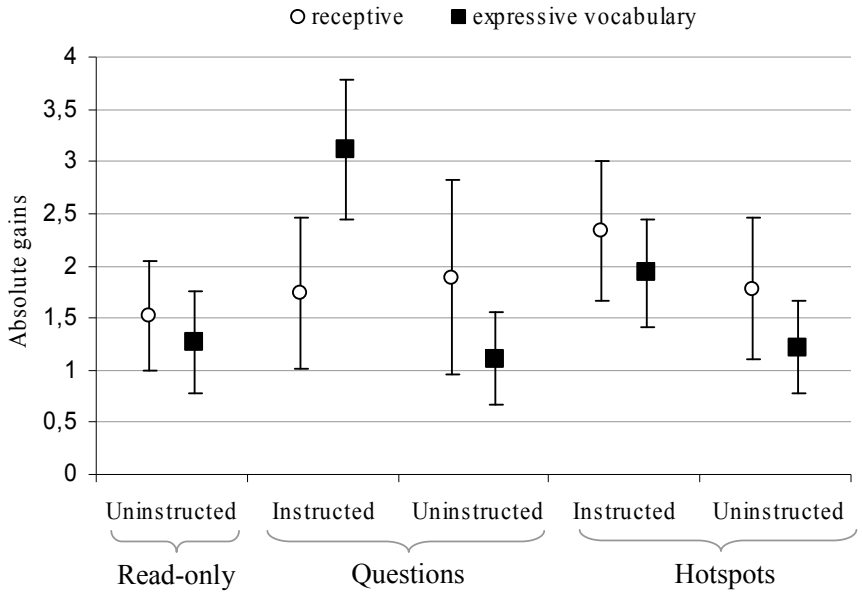

Figure 3.5. Receptive and expressive vocabulary gains in Experiment 2 (with confidence intervals), displayed per condition and word test. Gains are calculated by subtracting pretest from posttest scores with maximum scores of 8 per variable.

Effects on novel vocabulary

For analyses on words with zero scores on the pretests, we calculated the percentages of words that were learned receptively, expressively, both ways, and not at all. Because normality assumptions were not satisfactory for all variables, a nonparametric test (Wilcoxon) was preferred.

First, we examined how effective questions and hotspots were in promoting novel word learning. In each condition, about half of the target words were unknown at the receptive as well as expressive level at pretest (questions: 53% of instructed and 55% of uninstructed words; hotspots: 57% of instructed and 58% of uninstructed words). Figure 3.6 shows what percentage of instructed and uninstructed (novel) words was learned expressively, receptively, or both ways.

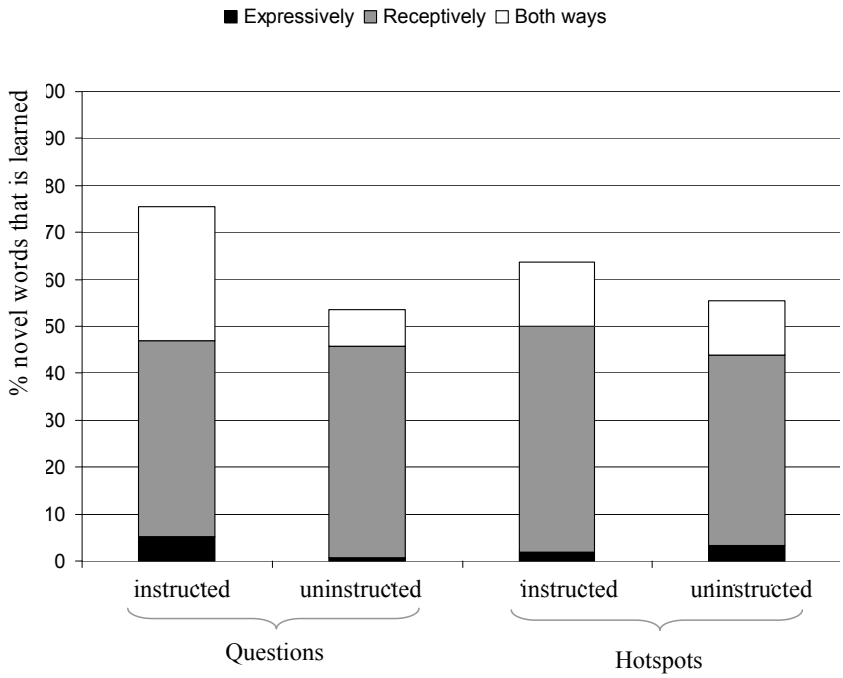

Figure 3.6. Percentage of instructed and uninstructed words unfamiliar at pretests that were learned expressively only (black), receptively only (grey), or both ways (white) in Experiment 2.

For both conditions separately, we contrasted gains in instructed and uninstructed words to examine how questions or hotspots affected word learning. MC questions advanced learning words expressively alone, $z = -2.02$, $p = .04$, $d = .85$, and learning words both ways, $z = -3.09$, $p = .002$, $d = 1.46$. An effect of questions was not found when words were learned receptively only, $p = .34$. Hotspots, by contrast, did not promote novel word learning beyond encounters in text, all p's $> .31$. A direct comparison between instructed words in the hotspots and in the question condition demonstrated that questions were more effective

than hotspots in advancing learning novel words both ways, z = -2.31, p = .02, d = .99. Thus, only with questions children could rapidly develop receptive and expressive knowledge about unknown words. There were, however, no significant differences between the two instruction methods in learning expressive, p = .13, or receptive knowledge alone, p = .36.

Effects on partially familiar vocabulary

Next we performed analyses on partially familiar words, i.e., passing the receptive but not the expressive pretest. Children learned more instructed than uninstructed words expressively as a result of MC questions, z = -2.51, p = .01, d = .48; differences approached significance in the hotspot condition, z = -1.75, p = .08, d = .34. A comparison of words learned expressively in the two experimental conditions did not reveal a statistically significant difference, p = .63, indicating that questions were as effective as hotspots in promoting expressive word knowledge for partly familiar words.

Repetition

The feedback procedure for MC questions could amount up to 4 repetitions of the same word if a child failed to correctly answer the question at the third attempt. Children who immediately select the correct image heard the target word only two times. To test whether more repetitions of target words might explain the effect of MC questions, we explored the relation between number of word repetitions in this condition and growth in vocabulary. In all, the average number of extra repetitions due to errors ranged between 0 and .88 per question. No significant correlations were found between number of repetitions and growth in either receptive, r = -.09, p = .71, or expressive vocabulary, r = .14, p = .55.

Discussion

Experiment 2 confirmed the finding of Experiment 1 in that MC questions in particular were beneficial for learning words expressively. Consistent with Experiment 1, a substantial percentage of instructed novel vocabulary (i.e., unknown at the receptive and expressive pretests) was learned receptively and expressively as a result of MC questions. Likewise, questions were beneficial for

learning words expressively when they were receptively known on pretest. There is no strong evidence that effects of questions result from differences in number of exposures to target words (e.g., De Temple &Snow, 2003). Correlations between the number of word repetitions in the questions condition and vocabulary learning were nonsignificant, indicating that word repetitions did not contribute much to growth in vocabulary. However, more controlled experiments on this matter are needed because one might argue that children who needed repetition were at lower levels of word understanding, thereby making this test invalid.

Results were less consistent for the effectiveness of hotspots as interactive devices because there were no effects of hotspots on completely novel words. On the other hand, hotspots were as effective as MC questions in enriching word knowledge of partly familiar words, even though the effect size for questions ($d = 0.48$) outperformed the effect size for hotspots ($d = 0.34$) and the result for hotspots was only marginally significant ($p < .08$). Thus, findings indicate that questions are more effective than hotspots, but the advantage depends on knowledge at the outset. Instruction format seems less important for words that are somewhat familiar.

Much to our surprise, Experiment 1's finding that receptive vocabulary improves as a result of questioning was not replicated. The best explanation for this anomalous outcome in Experiment 2 seems to be regression to the mean; the rather high mean pretest scores for instructed words in the questions condition (mean scores for questions and hotspots = 3.78 and 2.93, respectively) appear to move down on posttest relative to the population. However, listing all of the findings, the conclusion from Experiment 1 that MC questions are especially effective in promoting word knowledge still stands. This conclusion is consistent with studies of adult–child book sharing showing that extratextual vocabulary questions added to the reading are beneficial (Coyne, McCoach, Loftus, Zipoli, & Kapp, 2009; Ewers & Brownson, 1999; Sénéchal, 1997; Sénéchal et al., 1995). An alternative interactive device in e-books, hotspots, can also expand word knowledge, but only when words are receptively familiar.

Current findings suggest that the format of forcing children to reflect on small differences between words stimulates learning that goes beyond the acquisition of a label. An instructional method that requires children to label objects by selecting the correct image among alternatives (questions) was more effective than a method in which the label is provided without any effort of the child (hotspots). Thus, results are in line with the theory that increases in the amount

of invested mental effort give learning a boost (Neuman, 1997; Salomon, 1984). Choosing among various answers may make children reflect on fine-grained differences among words such as 'broken-hearted,' 'angry,' and 'shy,' which may stimulate more in-depth processing, thereby contributing to the acquisition of deeper meanings of words and fine-grained semantic differentiation (e.g., Coyne et al., 2009; McKeown et al., 2011; Nagy & Scott, 2000).

General discussion

Results of the current experiments show that electronic storybooks are most beneficial for word learning when they include extratextual vocabulary instructions. Questions were a valuable addition to other cues that can be derived from the visual and verbal context. However, familiarity with the target words was found to moderate learning with differential effects of instruction type on either novel or partly familiar words. When children have already acquired some receptive knowledge about the rare words in the story text, definitions (hotspots) and questions are equally effective; both methods appeared to be a boost for acquiring expressive knowledge. With some prior (receptive) knowledge, isolating and repeating the word seems to provide enough opportunities to expand further knowledge about the word's meaning and pronunciation (De Temple& Snow, 2003). However, when words are novel, as is indicated by zero scores on receptive and expressive pretests, only questions promoted expressive word learning beyond encounters in the story text; merely repeating and defining a word (hotspots) was not effective. Both devices, questions and hotspots, present words and pictures in a close temporal congruity, a practice that has been demonstrated to support memorizing and retaining words (Paivio, 1986). Our finding that the questioning technique is more effective than hotspots may indicate, nevertheless, that the key to instructing words effectively lies somewhere in the questions format.

There might be several explanations for the benefit of questions over definitions, and these explanations are not necessarily mutually exclusive. First of all, in the book reading paradigm, it is presumed that two-way interactive reading sessions that engage children are more effective in expanding vocabulary knowledge than one-way sessions in which children only 'receive' information (Ewers & Brownson, 1999; Sénéchal, 1997; Sénéchal et al., 1995). Likewise,

television programs that invite young children to participate actively (e.g., Blues Clues) have been found to support learning more than other educational programs (Crawley, Anderson, Wilder, Williams, & Santomero, 1999). Thus, it seems that children particularly benefit from interactive devices that encourage them to make meaning rather than take meaning (Moreno & Valdez, 2005).

Second, it seems important that children are stimulated to process word meanings more deeply, for instance, by reflecting on small differences between words (e.g., McKeown et al., 2011). Connectionist models of spoken word recognition propose that when retrieving the meaning of a word, there is competition between neighboring semantic and phonological representations in the network (e.g., Marslen-Wilson & Warren, 1994; Rodd, Gaskell, & Marslen-Wilson, 2004). Not only may MC questions contribute to strengthening the correct association, but also asking children to exclude incorrect alternatives may weaken incorrect associations. Hotspots lack the need to distinguish between the target and neighboring words, which may explain their much weaker effects on vocabulary growth.

Third, the nature of the instruction may be an important factor in its effectiveness. From a dynamic systems theory (Smith & Thelen, 1993), we might argue that it is not the amount of invested mental effort but rather the more instructive nature of MC questions that makes the difference; the question structure focuses children's attention and encourages assembly of visual and auditory information to favor meaning over other potential sources of interest (e.g., the rich detail of an illustration, the magnified details). The dynamics of the questioning setup supports meaning more intentionally and deliberately than the dynamics of a hotspot exposure that provides relevant information but does not force choice. The hotspot is in a sense more playful—giving children's attention more free rein—than the MCQ format, which is more coercive and, thus, more instructive.

Finally, a crucial component in the instructiveness of questions may be the presence of a computer assistant who provided feedback that immediately followed the child's response (Corbett & Anderson, 2001). Moreover, feedback addressed the child in a personal way, meaning that help is adjusted to characteristics of the user or to the user's interaction with the system (Vasilyeva, 2007). A main finding in another recent study of intelligent tutoring was that children's code-related skills increased as a result of the program, but only when the program included a computer tutor who gave personalized oral feedback to children's correct responses and errors (Kegel & Bus, 2011).

Limitations

Inescapably, the current study has limitations, among which in particular are the generalizations of our findings. First, children in the current sample scored relatively high on a standardized language test, which enabled them to benefit maximally from the verbal context. It would be interesting to examine whether questions are just as effective for children who are less advanced in word knowledge.

Second, the storybooks in the current experiments might not be a good representation of the e-book apps that are currently on the market. The sample of books used here was written by skilled and (often) award-winning authors, illustrated beautifully, and enriched with video effects and vocabulary instructions to support children's story comprehension. By contrast, the majority of stories that are available on the Internet do not exceed the format of static pictures accompanied by spoken text (Roskos, 2011). If interactive features are available, they often only include hotspots with funny animations that do not support the storylines.

Finally, we did not examine whether children acquired word knowledge beyond the story context and its particular illustrations. Therefore, we wonder whether conditions as described here add to semantic depth of word knowledge (e.g., Nation & Cocksey, 2009; Ouellette, 2006), also described as the amount of nuanced knowledge that is acquired about a word (Proctor, Uccelli, Dalton, & Snow, 2009). Critics may suggest that the expressive vocabulary test in the current experiments involves merely labeling an object. However, we presume that children needed more knowledge about a word for completing the sentences in this task because these often included contradictions (e.g., "This road is not narrow but. . . <wide>"), synonyms (e.g., "Pete is playing all alone. He is playing by. . .<himself>"), categorizing (e.g., "This bird is called a. . .<sea gull>"), or explanations of words (e.g., "Grandpa is sleeping for a little while. He is taking a. . . <nap>").

Some assume that vocabulary depth is best tapped by letting participants define words or provide synonyms because this expresses the extent of semantic representation (e.g., Nation & Snowling, 2004; Ouellette, 2006; Ouellette & Beers, 2010). However, defining words is confounded with kindergartners' inability to express themselves (Vermeer, 2001). We agree with Pearson, Hiebert, and Kamil's (2007) previous suggestion that expansion of the so far "grossly undernourished" (p. 282) field of vocabulary assessment is required.

Conclusions

The goal of the current experiments was to examine whether extratextual vocabulary instructions are a valuable addition to electronic storybook reading. By manipulating the timing of instructions (Experiment 1) and how much effort instructions require (Experiment 2), we gained some insight into the effectiveness of interactive devices. These findings may serve as guidelines in designing e-book apps but also help to specify how interactivity supports learning. First, extratextual instruction is an important addition to text exposure. Children learn more instructed difficult words than uninstructed difficult words. Second, how words are instructed seems important. A questioning format seems vital for acquiring knowledge for novel words. The finding that questions often outperform hotspots may suggest that active engagement is vital. However, it may be important to also test alternative explanations (e.g., the presence of a tutor) in further studies. Third, challenging children to respond by means of computer mouse clicking (rather than verbally) did not restrict growth in language production; on the contrary, it added to both receptive and expressive language, which implies great promise for the design of e-book apps. Finally, a limited number of interruptions in the flow of reading did not interfere with learning from exposure to the story text alone.