

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<https://hdl.handle.net/2066/214573>

Please be advised that this information was generated on 2020-11-24 and may be subject to change.



Moderating role of reading comprehension in children's word learning with context versus pictures

Sophie Gruhn¹ | Eliane Segers^{1,2} | Ludo Verhoeven^{1,3}

¹Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

²Instructional Technology, University of Twente, Enschede, The Netherlands

³Faculty of Arts, University of Curaçao, Willemstad, Curaçao

Correspondence

Sophie Gruhn, Behavioural Science Institute, Radboud University, Postbus 9104, 6500 HE Nijmegen, The Netherlands.
Email: c.gruhn@bsi.ru.nl

Funding information

Netherlands Initiative for Education Research (National Regieorgaan Onderwijsonderzoek) as a unit of the Netherlands Organisation for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek), Grant/Award Number: NWO 405-15-540-079

Abstract

We investigated pictorial versus contextual support effects over and above teaching of definitions on children's word learning and retention as well as the moderating role of reading comprehension. In a between-subject pre-post-retention test control group design, Dutch fourth graders learned concrete and abstract Dutch words. The context group received definitions and context sentences, the picture group received definitions and pictures, and the control group only received definitions. Children in the picture group outperformed children in the context group but not the control group on all words directly after learning. This learning gain, however, disappeared over time because participants in the context group retrieved the word meanings better than in the other groups 1 day later. Interestingly, reading comprehension moderated the forgetting effect, as low comprehenders in the experimental groups forgot less than those in the control group. Adding pictures to definitions, thus, leads to richer long-term knowledge for low comprehenders.

KEYWORDS

first language acquisition, multimedia, primary school, reading comprehension differences, vocabulary instruction

1 | INTRODUCTION

Vocabulary is an important predictor of reading comprehension abilities across primary school, and reading comprehension in turn is an important skill for vocabulary learning (e.g., Verhoeven, Van Leeuwe, & Vermeer, 2011). As reading comprehension is highly relevant for general future educational success, there is a high need of effective vocabulary instructions at school (Biemiller, 2006). Word learning during text reading becomes more efficient when word meanings are explicitly taught (Marulis & Neuman, 2010). Besides vocabulary breadth (quantity of words in the mental lexicon), specifically vocabulary depth (how much is known) is a significant and unique predictor of reading comprehension (e.g., Cain & Oakhill, 2014). To foster such

deeper word knowledge, extended direct vocabulary instruction with richer semantic information has been shown to be effective (e.g., Coyne, McCoach, Loftus, Zipoli, & Kapp, 2009). Such instruction may be especially helpful for children with lower reading comprehension abilities because they have problems in inferring word meanings from context during reading (Cain & Oakhill, 2011). To provide richer semantic information during word learning, often context sentences or pictures are added to definitions, albeit that the effectiveness of the latter is less clear for abstract words (e.g., National Reading Panel, 2000; Sadoski, 2005). So far, the unique benefits of context sentences or pictures over and above definitions on children's word learning have not yet directly been compared with word concreteness and reading comprehension differences being taken into account. In the

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. Journal of Computer Assisted Learning published by John Wiley & Sons Ltd

present study, we, therefore, investigated how adding context sentences or pictures to definitions supports Dutch fourth graders in their direct learning and retention of concrete and abstract words in relation to their reading comprehension.

Children learn new words by inferring their meanings from their daily input without explicitly being taught the meaning (Webb & Nation, 2017). For each word, a phonological, orthographical, and semantic representation is stored in memory, known as lexical representation (Perfetti, 2007, 2017). The quality of lexical representations is determined by the strength of each informational node and their interconnections. In terms of semantic representations, the quality increases with the degree to which new word meanings are integrated in the network of already existent and semantically related lexical representations (Read, 2004). Deep word knowledge contains all semantic features with which words from the same semantic field can be distinguished, includes information about the syntactical and morphological use together with its use in context of other words (Perfetti, 2007, 2017). The acquisition of deep semantic knowledge is an incremental long-term process, as frequent exposure and variation in context are relevant to construct successively deeper semantic knowledge (Webb & Nation, 2017). Without frequent and salient input along with active processing, children quickly forget newly encountered words (Vlach & Sandhofer, 2012). The complementary learning systems model of memory distinguishes two cognitive stages (Davis & Gaskell, 2009). Directly after the first encounter, the word representation is held as an episodic trace in the hippocampal system, which is bound to the specific occurrence and is yet isolated from the already existent knowledge. Next, a more generalized and abstract form is gradually integrated into long-term memory. Such lexical integration, also termed consolidation, requires time and is supported by repeated rehearsal or overnight sleep, as has been evidenced in children (van der Ven, Takashima, Segers, & Verhoeven, 2017).

Explicit teaching of word meanings during reading helps children to learn more words than if word meanings have to be inferred from context (Marulis & Neuman, 2010). A higher intensity of such instruction strengthens this effect not only in terms of the number of learned words but also in the acquired semantic depth and retention (e.g., Beck & McKeown, 2007; Coyne et al., 2009; Laufer & Rozovski-Roitblat, 2015). Instructions for vocabulary breadth rather provide a brief definition, whereas instructions focusing on vocabulary depth involve more frequent encounters in various contexts, activities with a deeper interaction with the word meanings and forms besides discussions about semantic features and the relationship to the already existent knowledge (e.g., Coyne et al., 2009). Examples are the rich vocabulary instruction approach of McKeown and Beck (2004) or the semantic feature analysis described by Pittelman, Heimlich, Berglund, and French (1991).

Providing definitions during explicit vocabulary instruction can be considered the default training form, which is often semantically enriched by adding context sentences and pictures to facilitate the development of deeper semantic knowledge (e.g., National Reading Panel, 2000; Sadoski, 2005). Due to the role of context in word learning, context sentences may naturally provide deeper semantic information

than definitions alone (Beck, McKeown, & Kucan, 2013). They are examples for the use of the new word in the context of other words and convey information on the co-occurrence of words as well as morphological and syntactical features (e.g., Bullinaria & Levy, 2007). Associated words will be activated in the mental lexicon while reading the context sentence, and this activation spreads to semantically related concepts (Collins & Loftus, 1975). This may help in linking the new word to the already existent knowledge and in recalling the information (McKenzie & Eichenbaum, 2011). Aligned with this, several studies showed that the combination of definitions and context sentences in explicit first language vocabulary instruction with children or college students led to better learning than providing definitions (Kolic, 1991; Stahl, 1983; Stahl & Fairbanks, 2006) or context sentences only (Bolger, Balass, Landen, & Perfetti, 2008). The former was also the case for children with low vocabulary skills for retention (Nash & Snowling, 2006). In contrast, this contextual benefit was not found in other studies with college students on first language learning (Smith, Stahl, & Neil, 1987) or with adults on foreign language learning (e.g., Golonka et al., 2015; Webb, 2007a).

Besides context sentences, adding pictures to definitions may improve children's word learning according to the cognitive theory of multimedia learning (Mayer, 2014). This theory is based on the idea of dual coding, which states that all knowledge is stored in memory via a verbal and nonverbal code (Paivio, 1991; Sadoski, 2005). Pictures may support mental imagery for word meanings and, therefore, can support word learning and retrieval when added to definitions. Within the cognitive theory of multimedia learning, it is assumed that visual and verbal (auditory) information are processed in different systems of the human working memory and that the amount of information which can be processed within one system is limited (Baddeley, 1999; Sweller, 1994). If information is perceived via text and pictures, the visual and verbal working memory channels can process different information parallel, which, thus, can lead to better and deeper learning by integrating the information from different sources than if information is presented by text or pictures only. This is known as the multimedia effect. Processing all information only via the visual or verbal system may lead to a cognitive overload and may disturb learning. According to the cognitive theory of multimedia learning, adding pictures to definitions may foster word learning more than adding context sentences to definitions. The benefit of multimedia implementation in instruction for children is widely supported (Bus & Neuman, 2009; Zhou & Yadav, 2017). Also, several studies on word learning with children or college students (Kim & Gilman, 2008; Smith et al., 1987) and adults (Plass, Chun, Mayer, & Leutner, 2003; Shen, 2010) provided evidence for a superior word learning effect, if verbal and visual information were combined during instruction of words compared with when only verbal or visual information was available. This benefit from multimedia use in word learning seems to account especially for children or adults with lower language abilities (Silverman & Hines, 2009; Yun, 2011). Not all studies showed, however, a benefit of adding pictures to verbal information during children's word learning (Acha, 2009; Cohen & Johnson, 2011). Pictures can also distract from storing the word form, as was observed for kindergarteners

practicing sight word reading of semantically familiar words (Elliott & Zhang, 1998). To truly enrich the semantic information of definitions, pictures need to provide information of additive value (Carney & Levin, 2002; Kalyuga & Sweller, 2014).

In the reviewed studies on the pictorial support effect, pictures were added to text in a combined form of definitions and context sentences. Furthermore, they were conducted with older children and adults or focused on word learning in a foreign language. The unique benefit of either adding context sentences or pictures to definitions has not yet directly been compared for young children who are learning new words in their first language. Based on the strong evidence for superior learning if text and pictures are combined than from text alone (Bus & Neuman, 2009; Mayer, 2014), adding pictures to definitions may lead to better learning than adding context sentences to definitions or from definitions alone.

The above mentioned studies have not taken individual and word level characteristics into account when they investigated contextual and pictorial support effects in word learning, although this is particularly relevant to develop efficient vocabulary instructions for struggling readers (Elleman, Steacy, Olinghouse, & Compton, 2017). It has been shown that higher reading comprehension abilities led to better word learning from context and definitions (e.g., Bolger et al., 2008; Cain, Oakhill, & Elbro, 2003; Cain, Oakhill, & Lemmon, 2004; Perfetti, Wlotko, & Hart, 2005). This link was explained by lower vocabulary along with lower integration and inference abilities of participants with problems in reading comprehension, which are highly relevant for word learning from context and definitions (Fukkink, 2005; Stafura & Perfetti, 2017). Children who struggle with reading comprehension lack strategic knowledge about how to make use of context and seem to construct less stable memory traces for newly encountered words (Nation & Snowling, 1999; Perfetti et al., 2005). Therefore, pictures might be especially useful to support word learning of low comprehenders.

In terms of word level differences, vocabulary learning with context or pictures may also depend on word concreteness because the representation of concrete and abstract words in the mental lexicon is conceptually different (Hoffman, 2016; Mestres-Missé, Münte, & Rodriguez-Fornells, 2014). Concrete meanings are directly linked to sensory experiences (e.g., shape, taste, and sound) and are based on specific semantic properties. In contrast, abstract meanings are difficult to imagine, are fuzzier, rely more on semantic associations, and vary more across contexts. The dual coding theory (Paivio, 1991; Sadoski, 2005) has specified that only concrete words are stored with a verbal and visual code, whereas abstract meanings are only represented verbally. Therefore, pictures are often not considered useful for abstract word learning. It has been shown, however, that pictures support learning of foreign abstract words in adults, whereas this was not the case for concrete words (Shen, 2010). For concrete words, adults were assumed to activate own mental pictures. Pictures would then have been redundant for concrete but not for abstract words (e.g., Kalyuga & Sweller, 2014). In terms of contextual support effect, stronger benefits may be assumed for learning of abstract words than of concrete words based on the context availability theory

(e.g., Schwanenflugel & Stowe, 1989). According to this theory, semantic (or contextual) information of abstract words is retrieved slower than of concrete words due to the larger variety of abstract word meanings in different context and due to the resulting wider spread of the semantic network. Concrete words are related to less semantic concepts, and their contextual information can directly be retrieved from the mental lexicon, which leads to a processing advantage (concreteness effect). This could be eliminated when abstract words were embedded in context sentences (e.g., Schwanenflugel & Stowe, 1989). The contextual benefit was even stronger for abstract than for concrete words in some studies (Hoffman, Jefferies, & Lambon Ralph, 2010), but this was not consistently found (Mestres-Missé et al., 2014).

In all, the unique benefit of context sentences or pictures over and above definitions on children's word learning is not yet clear, especially not when taking word concreteness and reading comprehension differences into consideration. In the present study, we investigated how adding context sentences or pictures to definitions supports primary school children in their first language direct word learning as well as retention of the knowledge for concrete and abstract words. We examined if reading comprehension functioned in these respects as a moderator. Dutch fourth graders learned a set of previously unknown Dutch concrete and abstract words. All participants read a written definition per word once. The context group also read a context sentence, and the picture group additionally received a picture. The control group only saw the word definition. The ability to define the words was measured directly after learning and 1 day later to assess consolidation.

Based on the cognitive theory of multimedia learning (Mayer, 2014), we first expected highest learning gains and retention for the picture group, followed by the context group, and then the control group. Second, we assumed that there were differences between concrete and abstract words across the groups driven by the theory that concrete knowledge is represented by a dual code (verbally and visually), whereas abstract knowledge is coded only verbally and can be facilitated more by contextual information (Paivio, 1991; Sadoski, 2005; Schwanenflugel & Stowe, 1989). Therefore, we hypothesized that the above mentioned pattern of group differences only holds for concrete words. For abstract words, we expected that learning gains and retention are higher in the context group than that in the picture and control group due to the difficulty to represent abstract meanings with pictures, whereas the picture and control group do not differ from each other. Participants with lower reading comprehension abilities are considered to have problems in making use of verbal information in word learning such as context sentences and definitions (e.g., Bolger et al., 2008; Cain et al., 2004). In contrast, information from pictures can directly be retrieved without reading. Therefore, our third hypothesis was that towards lower levels of reading comprehension, the pictorial support effects become stronger, whereas the contextual support effects degrade. In other words, children with lower reading comprehension abilities were hypothesized to perform best in the picture group while performing similar in the context and control group.

2 | METHOD

2.1 | Participants

This study was conducted in spring 2017 with 191 Dutch-speaking fourth graders from seven different schools (nine classes) in urban and suburban south-eastern regions of the Netherlands. Two participants were excluded because no data were available from the experiment or on reading comprehension. The final sample of 189 participants had a mean age of 9 years and 10 months ($SD = 5.58$ months). Participants were allocated randomly within classrooms to one experimental group: context group (males = 32, females = 31), picture group (males = 28, females = 38), or control group (males = 30, females = 30). Teachers indicated for some participants, use of another or additional home language besides Dutch ($n = 10$), diagnosed dyslexia, or other language related ($n = 17$) and developmental problems ($n = 8$). Those participants were equally spread over the groups and were not excluded from the analysis because this reflects an authentic classroom in the Netherlands and as such provides ecological validity to our study. Passive parental consent was received. Ethical approval was obtained from the Ethics Committee of the Faculty of Social Sciences of our university with the number ECSW2016-2811-448.

2.2 | Materials

2.2.1 | Items

We tested the knowledge of fifth graders on a set of infrequent nouns from a Dutch corpus (Tellings, Hulsbosch, Vermeer, & Bosch, 2014) to choose words that were most likely unfamiliar to fourth graders in our study. This was supplemented by words from an earlier study (van der Ven et al., 2017). The final set consisted of 10 concrete (e.g., fortress) and 10 abstract nouns (e.g., charisma) with a low written lemma frequency ($M = 3.45$, $SD = 5.60$, range 0–17, Tellings et al., 2014). Word concreteness was rated dichotomously by two native-speaking authors of this study based on the ease to represent words with pictures, as concreteness and imageability are highly correlated (Connell & Lynott, 2012). The concreteness scores retrieved from a large Dutch database (Brysbaert, Stevens, De Deyne, Voorspoels, & Storms, 2014) were significantly higher for concrete than for abstract words, $t(13) = -7.33$, $p < .001$. Five concrete words were not listed in this database because we aimed to select very unfamiliar words based on van der Ven et al. (2017). Per word, a written definition was formulated. Additionally, a context sentence and a picture were designed to support the definition (see Appendix A). The context sentence aimed to provide a natural context at sentence level. The definition and context sentences were based on information provided in an online dictionary (Van Dale, 2017) or a lexicon for children (Winkler Prins, 2012). The length of the definitions ($M = 6.45$ words, $SD = 0.88$ words) and context sentences ($M = 9.00$ words, $SD = 1.07$ words) were kept equal between items. The pictures were chosen

from free online picture sources, the website DigiWak (Stichting Digiwak, 2017), or a child lexicon (Winkler Prins, 2012).

2.2.2 | Vocabulary experiment

Within one session, children studied 20 Dutch words (10 concrete and 10 abstract) with help of the definitions. The meaning was provided only once. The experiment was conducted on a Dell Latitude E5450 laptop with a 13.2-in. screen. The experiment was developed with Psychopy and Python. The defined word appeared in the upper left hand corner on a grey screen in 36-point MS Shell Dlg 2 font. Below the word, the definition was printed in 26-point of the same font coloured in blue. Children could continue to the next item after 5 s, when an arrow button appeared in the bottom right-hand corner.

For the context group, 2 s after the definition was provided, a green-coloured context sentence was presented below the definition in the same font and size. For the picture group, 2 s after the definition was provided, a picture (800 × 600 pixels) appeared below the definition. In the control group, no additional semantic information was provided. Figure 1 provides a translated example per group.

The experimenter instructed the children to learn and remember the meaning of the words because they would be asked to provide it afterwards in their own words. The experiment was preceded by a practice item for a well-known word. The children could go through the words on a self-paced speed and were made aware that it was not possible to go backwards for restudying. The words were taught in two steps of each 10 words with an intermediate break. During the break, children conducted individual tasks to avoid interim discussions about the word meanings. To avoid difficulty differences between the words of each step, the items were separated in two comparable blocks by controlling for concreteness, written lemma frequency, orthographic regularity, and syllable number. The presentation of blocks was counterbalanced across the two steps of the experiment (93 participants started the experiment with word Block 1 and 96 participants with word Block 2). The words within blocks were presented in one of 50 random orders.

2.3 | Procedure

In the first session, testing on the control measures and the prior semantic knowledge on the target words took place individually for approximately 20 min. In the second session, participants studied 20 Dutch word meanings via a computerized experiment in groups of maximally eight children. To reduce the memory load, the words were taught in two steps of 10 words each with a short intermediate break. Directly after each step, the acquired semantic knowledge was tested (posttest). All children were tested again in class 1 day after the experiment (retention test). The second and third session took about 20 min, respectively. Each session was conducted

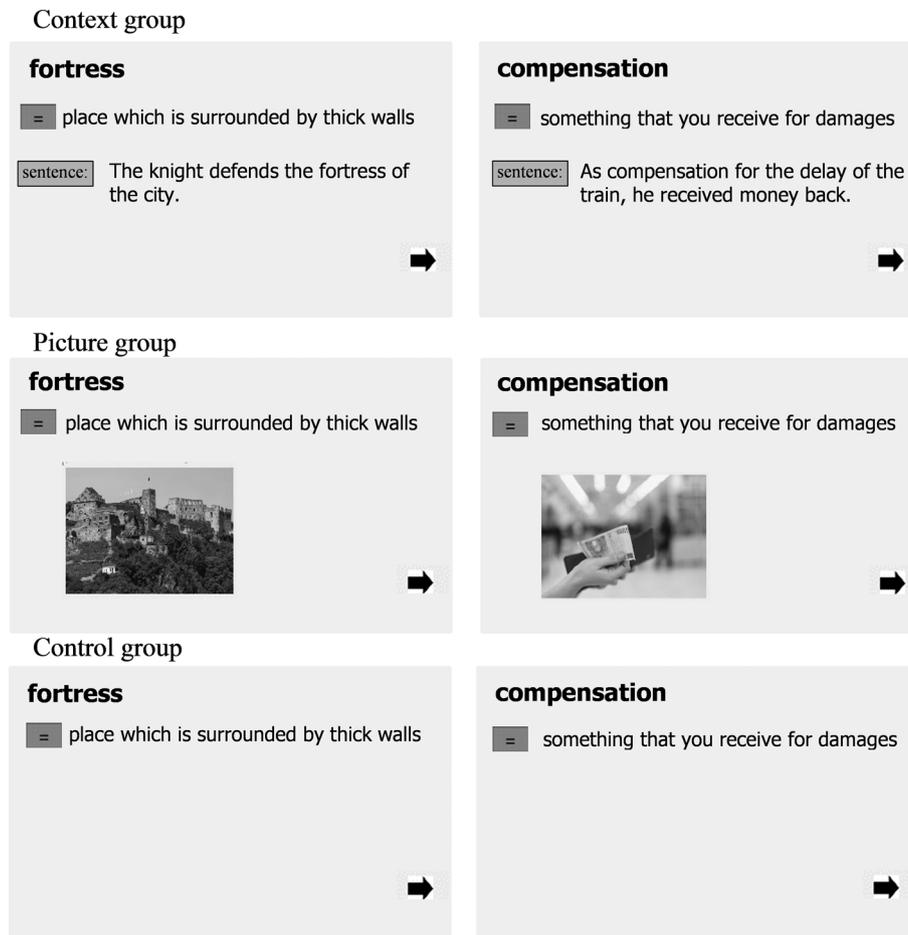


FIGURE 1 To English translated example of the semantic information provided during the vocabulary experiment in each group for a concrete word (left) and an abstract word (right). The original font size and distances between units as presented on the computer screen were adjusted for this illustration.

on a separate day within maximally 2 weeks. Data were collected by three experimenters.

2.4 | Measures

2.4.1 | Control measures

Visual memory

The Corsi block-tapping task, as described in Kessels, Van Zandvoort, Postma, Kapelle, and De Haan (2000), was conducted to assess visual memory. The experimenter tapped on wooden blocks in a certain order with an increasing number of blocks ranging from two to nine. Subsequently, the participant had to repeat this sequence. Per sequence length, two trials were conducted. The product of the maximum number of blocks repeated correctly, and the total number of correct trials resembled the visual memory capacity. Cronbach's alpha for this task was .75 in a study conducted with 25 neurocognitive disordered participants (De Paula, Malloy-Diniz, & Romano-Silva, 2016) and .61 in a study conducted with a larger and healthy sample under slightly different task conditions (Busch, Farrell, Lisdahl-Medina, & Krikorian, 2005).

Verbal memory

A digit span task, taken from the Dutch translation of the Wechsler Intelligence Scale for Children (WISC-III NL, Kort et al., 2005) was used to measure verbal memory. Participants had to repeat auditory presented rows of numbers with increasing length in the same or reversed order. Per digit span, ranging from two to nine, two trials were conducted. The test was stopped after two incorrect trials of the same length. Cronbach's alpha of the task was on average .64, which indicated a low but acceptable reliability considering the low number of items (Kort et al., 2005; Pallant, 2013).

Reading fluency

An isolated Dutch word reading task (1-min test, Brus & Voeten, 1979, Form B) was used to indicate reading fluency. Children had to read words of different length from a list as quick and accurate as possible. The test score represented the total number of words that were read correctly within 1 min. As the nature of the test does not allow calculating values of internal reliability, Brus and Voeten (1979) provided alternative measures. The correlation between parallel test forms (Forms A and B) at the same measurement moment was .92 to .94,

and the test-retest reliability was .85 to .87 for the same form or parallel forms.

Vocabulary

The schools provided us with vocabulary scores of a nationally standardized test for 122 participants (Reading Vocabulary; Verhoeven & Vermeer, 1999). Those tests are used by 95% of Dutch schools to monitor learning progress (Cito, 2018a, 2018b). We conducted for the remaining 67 children a standardized vocabulary subtest from the Dutch test battery Language Test for Allochthone Children (Reading Vocabulary; Verhoeven & Vermeer, 1993). In both tests, children had to choose from four options the correct definition or synonym for a target word, which was embedded in a sentence. The tests consisted of the same items but were scored on a different scale. The predicted passive vocabulary breadth scores from the test manuals, therefore, were used as a common vocabulary measure for the whole sample. Cronbach's alpha for both tests was .87 (Verhoeven & Vermeer, 1996).

2.4.2 | Reading comprehension

We received from schools reading comprehension scores from nationally standardized tests (Feenstra, Kleintjes, Kamphuis, & Krom, 2010; Staphorsius & Krom, 1998). In those tests, children had to read short texts and to answer multiple choice questions. The scores derived from two different test editions widely used in the Netherlands for monitoring of student's learning progress (Cito, 2018a, 2018b). Their scores were comparable because they were based on the same ability scale with item response theory (Tomesen, Weekers, Hiddink, & Jolink, 2017). For both test versions, *M*Acc was higher or equal to .89 (Staphorsius, Krom, Kleintjes, & Verhelst, 2004; Tomesen et al., 2017).

2.4.3 | Semantic knowledge

At pretest, we checked the prior knowledge of children on the items from the experiment. The experimenter read the words aloud in a speed of one word per second, while the participants listened and silently joined reading. By this, children were familiarized with the phonological word form and influences of decoding problems were kept to a minimum. If participants knew a word meaning, they rang a bell, explained the meaning, and wrote it down on paper. Beforehand, a practice item was conducted to explain how words can be defined. At posttest, directly after each word block, children wrote down definitions on a sheet of paper where all words forms of the respective word block were listed besides a practice item. The same procedure was followed at retention test, except that words were not separately presented in blocks. At no point, children received feedback. At all measurements, the 20 words were provided in one of 20 different orders.

Following Verhoeven and Vermeer (1993), the semantic knowledge was scored with 2 (*good knowledge*), if the definition provided all meaning elements or a synonym. A score of 1 (*some knowledge*) was given, if the child provided examples, characterized or categorized the word, or if only some of the meaning elements were named. All

other cases were scored with 0 (*no knowledge*). For each word, a list of meaning elements and examples for the scoring of answer options was provided to the raters. The reliability of the pretest was low (Cronbach's $\alpha = .43$), which was expected because most children did not have prior knowledge. Cronbach's alpha for the posttest was .75 and for the retention test .70, which is considered an acceptable reliability (Pallant, 2013). To establish an interrater reliability measure, the three experimenters rated independently from each other the provided word definitions at posttest for a subsample of 40 participants. The interrater reliability was considered high because the inter-item correlation between the raters ranged from .90 to .93 and the intraclass correlation was .97 (Field, 2009).

2.5 | Data analysis

We used the sum scores of the semantic knowledge at each measurement for the analyses. As they deviated from the normal distribution (skewness or kurtosis divided by their standard errors extending the absolute value of 2), we used square root transformed sum scores for the analyses (Field, 2009; Tabachnik & Fidell, 2013). After the transformation, the posttest and retention test scores were normally distributed, but this could not be reached for the pretest scores. That some children knew a few words at pretest, whereas most children did not, explained the strong deviation. We, therefore, decided to use gain scores from pretest to posttest in the analyses of learning gains (normally or close to normally distributed), to control for prior knowledge. The change in the semantic knowledge from posttest to retention test was analysed separately as forgetting effect. To answer each research question separately, we conducted several analyses steps with the relevant variables taken into account, as described below. When the moderational role of reading comprehension was investigated, reading fluency was added as covariate to control for decoding abilities. If not further specified, assumptions of the statistical analyses were met. Reading comprehension and reading fluency were centred for the analyses. The analyses were conducted in SPSS Version 23, and figures were created in R with the package *jtools* (Long, 2018).

3 | RESULTS

3.1 | Descriptive statistics and group differences on the control variables

The participants in the context, picture, and control group did not differ in age, control measures, and reading comprehension. An overview about the participant information and results on the group comparisons are provided in Appendix B. The groups did not differ significantly on their semantic knowledge at pretest (Kruskal-Wallis tests for all words and separately for concrete and abstract words: $\chi^2_s(2) = 0.61$ to 1.80, $ps = .41$ to .74). Table 1 presents an overview of the mean sum scores of the semantic knowledge.

TABLE 1 Semantic knowledge: Mean scores per measurement, group, and word concreteness

Measurement and type of words	Group							
	Total (n = 189)		Context group (n = 63)		Picture group (n = 66)		Control group (n = 60)	
	M	SD	M	SD	M	SD	M	SD
Pretest								
All words	0.51	1.16	0.54	1.13	0.39	0.93	0.60	1.40
Concrete words	0.29	0.83	0.30	0.75	0.24	0.68	0.33	1.04
Abstract words	0.22	0.61	0.24	0.56	0.15	0.50	0.27	0.76
Posttest								
All words	7.88	6.05	7.02	6.44	8.79	5.85	7.78	5.81
Concrete words	5.11	3.79	4.68	4.01	5.50	3.75	5.13	3.60
Abstract words	2.77	2.94	2.33	3.08	3.29	2.90	2.65	2.78
Retention test								
All words	5.24	4.79	5.11	4.94	5.52	4.54	5.07	4.96
Concrete words	3.49	3.11	3.54	3.28	3.67	3.00	3.23	3.08
Abstract words	1.75	2.34	1.57	2.28	1.85	2.11	1.83	2.66

Note. *n* items = 20, containing 10 abstract words and 10 concrete words.

3.2 | Research Question 1: Group differences in the learning gain and forgetting effect

To investigate group differences in the learning gain, we conducted a one-way analysis of variance (ANOVA) with the gain scores of all words and group as between-subject factor (context, picture, and control). The results are presented in Table 2. The groups differed in learning gain, as there was a significant main—albeit small—effect of group. Bonferroni corrected post hoc tests revealed that participants in the picture group had higher learning gains than participants in the context group (see Appendix C, Figure C1). Other group differences were not significant.

Second, we examined the forgetting effect of all words with a repeated measures ANOVA with time as within-subject factor

(posttest and retention test) and group as between-subject factor (context, picture, and control). The forgetting effect did not differ between the groups, as the interaction effect between time and group was not significant (see Table 2).

3.3 | Research Question 2: Differences in learning concrete or abstract words

A repeated measures ANOVA was performed with the gain scores of concrete and abstract words to further specify how the previously found main effect of group for the learning gain of all words varied across word concreteness. Group functioned as between-subject factor (context, picture, and control) and word concreteness as within-subject factor (concrete and abstract). The learning gain did not differ between concrete and abstract words as the interaction effect between group and word concreteness was not significant. The main effect of group and respective group differences remained the same. The results are presented in Table 3.

To investigate group differences in the forgetting effect for concrete and abstract words, we conducted a repeated measures ANOVA with time (posttest and retention test) and word concreteness (concrete and abstract) as within-subject factors. Group was used as between-subject factor (context, picture, and control). The forgetting effect of concrete or abstract words did not differ between the groups as the three-way interaction effect between time, group, and word concreteness was not significant (see Table 3). The interaction effect between time and group became significant when taking word concreteness into account.

To better understand the significant interaction effect between time and group after taking word concreteness into account, we tested the interaction effect between time and group for each group contrast separately. This revealed that participants in the context group retained the semantic knowledge better than participants in the picture group and control group. Those were small to medium effects ($\eta^2_p = .05$ to $.07$). No differences were observed between participants in the picture and control group. Levene's tests for homogeneity of variance were significant for the comparison of the posttest

TABLE 2 Research Question 1: Learning gain and forgetting effect of all words

Variables and tests	Test statistic (df1, df2)	<i>p</i>	Effect size	Significance of the group differences		
				Context vs control group	Picture vs. control group	Context vs. picture group
Learning gain: one-way ANOVA with gain scores						
All words	3.53 (2, 186)	.03	$r = .19$.54	.63	.026
Forgetting effect: repeated measures ANOVA						
Time	129.85 (1, 186)	<.001	$\eta^2_p = .41$	n.a.	n.a.	n.a.
Group	1.22 (2, 186)	.30	$\eta^2_p = .01$	1.0	1.0	.36
Time × Group	2.84 (2, 186)	.061	$\eta^2_p = .03$	n.a.	n.a.	n.a.

Note. Gain scores are the difference between pretest and posttest scores. The forgetting effect is the change of the semantic knowledge from posttest to retention test. Levels of time: posttest and retention test; levels of group: context group, picture group, and control group. The significance of group contrasts was tested with Bonferroni corrected post hoc tests for main effects and with follow-up tests per group contrast for interaction effects. n.a., post hoc tests not available or not applied when the main or interaction effects were not significant.

TABLE 3 Research Question 2: Learning gain and forgetting effect of concrete and abstract words

Variables and tests	Test statistic (df1, df2)	p	Effect size	Significance of the group differences		
				Context vs. control group	Picture vs. control group	Context vs. picture group
Learning gain: repeated measures ANOVA with gain scores						
Group	4.47 (2, 186)	.01	$\eta^2_p = .05$.28	.66	.01
Word concreteness	93.39 (1, 186)	<.001	$\eta^2_p = .33$	n.a.	n.a.	n.a.
Group \times Word concreteness	0.93 (2, 186)	.40	$\eta^2_p = .01$	n.a.	n.a.	n.a.
Forgetting effect: repeated measures ANOVA						
Time	136.88 (1, 186)	<.001	$\eta^2_p = .42$	n.a.	n.a.	n.a.
Group	1.57 (2, 186)	.21	$\eta^2_p = .02$	1.0	1.0	.24
Word concreteness	113.14 (1, 186)	<.001	$\eta^2_p = .38$	n.a.	n.a.	n.a.
Time \times Group	4.55 (2, 186)	.012	$\eta^2_p = .05$.004	.83	.016
Time \times Word concreteness	1.48 (1, 186)	.23	$\eta^2_p = .01$	n.a.	n.a.	n.a.
Group \times Word concreteness	0.32 (2, 186)	.72	$\eta^2_p = .003$	n.a.	n.a.	n.a.
Time \times Group \times Word concreteness	0.84 (2, 186)	.44	$\eta^2_p = .01$	n.a.	n.a.	n.a.

Note. Gain scores are the difference between pretest and posttest scores. The forgetting effect is the change of the semantic knowledge from posttest to retention test. Levels of time: posttest and retention test; levels of group: context group, picture group, and control group; levels of word concreteness: concrete words and abstract words. The significance of group contrasts was tested with Bonferroni corrected post hoc tests for main effects and with follow-up tests per group contrast for interaction effects. n.a., post hoc tests not available or not applied when the main or interaction effects were not significant.

scores between context and control group (concrete and abstract words: $F_s(1, 121) = 4.78$ to 4.94 , $ps = .03$).

3.4 | Research Question 3: Moderating role of reading comprehension

Finally, we examined the moderating role of reading comprehension for the learning gain and forgetting effect of the total semantic knowledge by controlling for reading fluency. First, we conducted a univariate general linear model for the gain scores of all words with group as between-subject factor (context, picture, and control). We added reading comprehension as a main effect and the interaction between group and reading comprehension to the model. The results are presented in Table 4. Reading comprehension did not moderate the learning gain, as the interaction effect between group and reading comprehension was not significant. The main effect of group and respective group differences remained the same.

Second, to examine the forgetting effect, we conducted a repeated measures ANOVA with time as within-subject factor (posttest and retention test) and group as between-subject factor (context, picture, and control). We tested specifically for the interaction effect between time, group, and reading comprehension and entered the underlying main and two-way interaction effects to the model. Reading comprehension moderated the forgetting effect, as the interaction effect between time, group, and reading comprehension was significant (see Table 4).

To better understand the moderating role of reading comprehension for the forgetting effect, we examined the interaction effect

between time and group (context, picture, and control) separately for each of the 33% extreme performing reading comprehension groups (33% with lowest reading comprehension test scores and 33% with highest reading comprehension test scores) with otherwise only the underlying main effects and reading fluency in the model. A similar approach was also followed by Plass et al. (2003). We considered participants with reading comprehension scores smaller or equal to 159 as low comprehenders (context group $n = 20$, picture group $n = 19$, and control group $n = 22$) and participants with scores higher or equal to 181 as good comprehenders (context group $n = 20$, picture group $n = 21$, and control group $n = 28$). The interaction effect between time and group was only significant for low comprehenders. Figure 2 shows for low and good comprehenders how the forgetting effect differed between the groups. To further understand the significant interaction effect between time and group for low comprehenders, follow-up tests were conducted for each group contrast (context vs. control, picture vs. control, and context vs. picture). Low comprehenders in the control group forgot more than low comprehenders in the context group and more than their peers in the picture group. The group differences were medium to large ($\eta^2_p = .11$ to $.15$). The forgetting effect did not differ for low comprehenders in the context and picture group.

4 | DISCUSSION

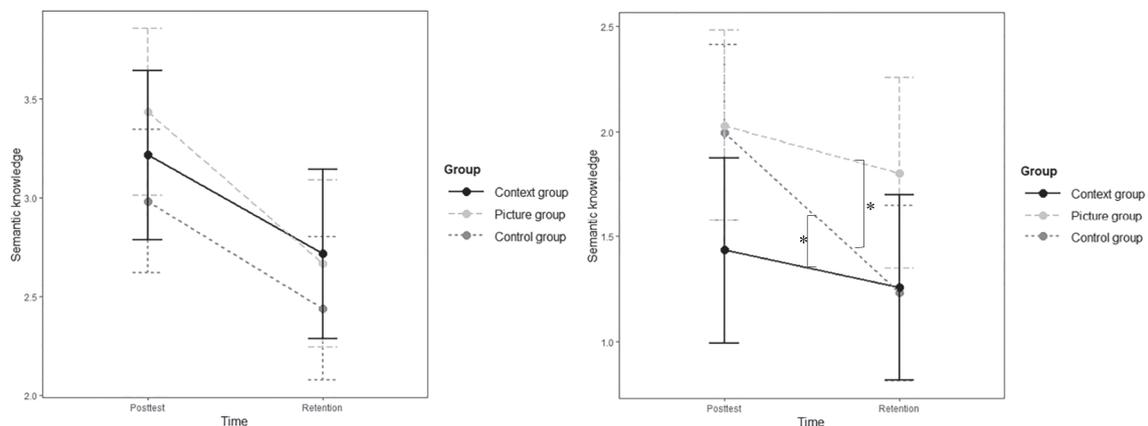
This study investigated how adding pictures or context sentences to definitions during explicit teaching of concrete and abstract word meanings benefitted the direct word learning and retention of Dutch

TABLE 4 Research Question 3: Moderational role of reading comprehension for all words

Variables and tests	Test statistic (df1, df2)	p	Effect size	Significance of the group differences		
				Context vs. control group	Picture vs. control group	Context vs. picture group
Learning gain: univariate general linear model for gain scores of all words						
Group	3.36 (2, 182)	.04	$\eta^2_p = .04$.67	.57	.03
RF	5.37 (1, 182)	.02	$\eta^2_p = .03$	n.a.	n.a.	n.a.
RC	12.72 (1, 182)	<.001	$\eta^2_p = .07$	n.a.	n.a.	n.a.
RC × Group	0.30 (2, 182)	.74	$\eta^2_p = .003$	n.a.	n.a.	n.a.
Forgetting effect: repeated measures ANOVA						
RF	19.59 (1, 182)	<.001	$\eta^2_p = .10$	n.a.	n.a.	n.a.
Time	135.21 (1, 182)	<.001	$\eta^2_p = .43$	n.a.	n.a.	n.a.
Group	1.24 (2, 182)	.29	$\eta^2_p = .01$	1.0	.91	.38
RC	31.62 (1, 182)	<.001	$\eta^2_p = .15$	n.a.	n.a.	n.a.
Time × RF	1.07 (2, 182)	.30	$\eta^2_p = .01$	n.a.	n.a.	n.a.
Time × Group	2.73 (2, 182)	.07	$\eta^2_p = .03$	n.a.	n.a.	n.a.
Time × RC	7.38 (1, 182)	.01	$\eta^2_p = .04$	n.a.	n.a.	n.a.
Group × RC	0.88 (2, 182)	.42	$\eta^2_p = .01$	n.a.	n.a.	n.a.
Time × Group × RC	4.55 (2, 182)	.012	$\eta^2_p = .05$	n.a.	n.a.	n.a.
Follow-up analysis						
Low RC Time × Group	3.95 (2, 57)	.025	$\eta^2_p = .12$.013	.039	.76
Good RC Time × Group	0.94 (2, 65)	.40	$\eta^2_p = .03$	n.a.	n.a.	n.a.

Note. Gain scores are the difference between pretest and posttest scores. The forgetting effect is the change of the semantic knowledge from posttest to retention test. Levels of time: posttest and retention test; levels of group: context, picture, and control. Reading comprehension and reading fluency were continuous variables and were centred for the analyses. The significance of group contrasts was tested with Bonferroni corrected post hoc tests for main effects and with follow-up tests for interaction effects per group contrast. n.a., post hoc tests not available or not applied when the main or interaction effects were not significant.

Abbreviations: RC, reading comprehension; RF, reading fluency. Low RC, low comprehenders (context group $n = 20$, picture group $n = 19$, and control group $n = 22$) with reading comprehension scores smaller or equal to 159. Good RC, good comprehenders (context group $n = 20$, picture group $n = 21$, and control group $n = 28$) with reading comprehension scores higher or equal to 181.

**FIGURE 2** Mean sum scores of the semantic knowledge at posttest and retention test (square root transformed) differentiated for participants with good reading comprehension (left) and low reading comprehension (right). * $p < .05$

fourth graders. We further examined the role of reading comprehension as a moderator of these effects.

The picture group showed higher learning gains from pretest to posttest than the context group, but both experimental groups did not

differ from the control group. The pictorial support effect disappeared over time, as the groups did not differ significantly anymore 1 day after learning. When word concreteness was taken into account, it was revealed that participants in the context group retained the meaning

of both concrete and abstract words better from posttest to retention test than participants in the picture and control group. Reading comprehension did not affect the learning gain but did impact the forgetting effect. Both the pictorial and contextual support effect for the knowledge retention became stronger towards lower levels of reading comprehension as compared with the control group. No differences were observed for children with higher reading comprehension abilities.

4.1 | Pictorial support effects

That adding pictures to definitions in contrast to adding context sentences to definitions benefitted the direct learning gain of all learners can be explained by the cognitive theory of multimedia learning (Mayer, 2014). Aligned with this theory and the descriptive results, it could be speculated that context sentences tended to have a distracting effect, which would, however, require further research. Similar pictorial support effects were reported by Smith et al. (1987) and by Kim and Gilman (2008) for word learning when pictures were added to definitions and context sentences.

In contrast to the studies of Smith et al. (1987) and Kim and Gilman (2008), the pictorial support effect did not sustain over time. This might be related to the low teaching intensity because the children only went through the definitions once, whereas restudying was possible in the studies of Smith et al. (1987) and Kim and Gilman (2008). It has also been shown that overconfidence in the mnemonic power of pictures may deteriorate pictorial support effects, as undergraduate students only learned new foreign words better with pictures than translations, when this overconfidence was eliminated (Carpenter & Olson, 2012). Furthermore, the studies of Smith et al. (1987) and Kim and Gilman (2008) were conducted with older participants or on foreign language learning. Older participants may differ in their learning abilities, memory, and retrieval capacities from our younger participants because their working memory system is still under development (Baddeley, Gathercole, & Papagno, 1998; Mann, Newhouse, Pagram, Campbell, & Schulz, 2002). It is also widely supported that foreign or second language learning is not comparable with first language learning (e.g., Choi, 2016; Meisel, 2011).

The finding that the pictorial support effect sustained over time for participants with low reading comprehension abilities but not with good reading comprehension abilities in terms of the retention of the knowledge is in line with previous studies that concluded that multimedia instruction is especially useful for low ability learners (Silverman & Hines, 2009; Yun, 2011). Good comprehenders might already activate broader semantic information when reading the definitions, and by this, pictures probably did not provide sufficient additional semantic information (Kalyuga & Sweller, 2014). Thus, the effect may disappear quicker over time. Our results did not support the assumption of Plass et al. (2003) that learners with lower verbal abilities lack cognitive resources to integrate the visual and verbal information. In contrast to the task used by Plass et al. (2003), a complex integration of verbal information during word learning was not strictly required in our self-paced experiment (Mayer, 2014; Sweller, 1994).

4.2 | Contextual support effect

Although there is plenty of evidence for contextual support effects on children's word learning (e.g., Stahl, 1983; Stahl & Fairbanks, 2006), context effects in our study did not reach significance for the direct word learning and retention effect. It has been shown that more repetitions or exposure to various context sentences resulted in better word learning (Bolger et al., 2008; Stahl & Fairbanks, 2006; Webb, 2007b) and that context effects may just become evident on long term (Nash & Snowling, 2006). Other explanations could be that the context sentences did not sufficiently enrich the semantic information of definitions or that the participants did not actively process the information. Because the definitions of the words were simultaneously provided with the context sentences, meaning inferences were not required. Previous studies that found a context effect required inferences from context or involved other meaning related activities (e.g., Coyne et al., 2009; Nash & Snowling, 2006; Stahl, 1983). Also, Bolger et al. (2008) found that the variability of context sentences did not matter for word learning when definitions were simultaneously available.

We found a contextual support effect for participants with low reading comprehension abilities, despite their problems in making use of contextual information in word learning after controlling for decoding abilities. This, however, only accounted for the knowledge retention and not for the direct learning effect. Therefore, context sentences might have supported low comprehenders rather in the lexical integration or retrieval than in the enrichment of provided semantic information. Contextual information spreads the activation to related semantic concepts, which may foster consolidation (Collins & Loftus, 1975; McKenzie & Eichenbaum, 2011). Because consolidation is considered to occur overnight, the effect might have become visible only 1 day later (van der Ven et al., 2017). The acquired lexical representations from learning with definitions only were probably less stable for low comprehenders. Similar findings were reported by Nash and Snowling (2006). Related to the complementary learning systems model (Davis & Gaskell, 2009), low comprehenders might also have difficulties in integrating a more generalized representation into their mental lexicon, and therefore, context sentences may have facilitated retrieval. Their representations might be in general fuzzier, which is why additional contextual information helps in processing, as is assumed for abstract words in the context availability theory (Schwanenflugel & Stowe, 1989). It could also be speculated that low comprehenders used the context sentences as a mnemonic for retrieval by inherently imagining pictures resembling the context sentences during the learning phase. Within a word learning strategy, the so-called keyword strategy, pictures, or the creation of mental images served as mnemonic by representing the link between the meaning of new words and an already familiar, similar sounding word (e.g., Mastropieri, Scruggs, & Mushinski Fulk, 1990). Previous studies showed that children were able to imagine pictures for word meanings via context sentences (Cohen & Johnson, 2011; Elliott & Zhang, 1998). They were, however, receiving instructions.

It remains unclear why context sentences enhanced neither the direct learning nor the retention effect of good comprehenders. Possibly, they might activate related words automatically on their own by reading the definitions and did not consider context sentences as relevant additional information, which could have covered the contextual support effect (Kalyuga & Sweller, 2014). Previous research also showed that richer contextual information makes it easier to guess the word meaning during learning, which may cause less thorough processing and by this worse retention (e.g., Bjork & Kroll, 2015; Golonka et al., 2015; Mondria & Wit-de Boer, 1991).

4.3 | Effect of word concreteness

Taking word concreteness into account further specified the results. First, the pictorial support effect for direct word learning seemed to be valid for both concrete and abstract words. Second, there was a general contextual support effect for the retention of the meanings of all words, as the context group forgot less from posttest to retention test than participants in the picture or control group.

That the pictorial support effect did not differ between concrete and abstract words in the direct learning gain contradicts the assumption that concrete words are better taught with pictures and abstract meanings better verbally (e.g., Sadoski, 2005). Other studies with children who used mixed types of words also found a benefit of pictures in word learning (Kim & Gilman, 2008; Smith et al., 1987). Aligned with the findings of Shen (2010) for adults, the fact that there is a similar benefit for abstract words may result from the additional visual support, which is naturally not evoked when implementing new lexical representations for abstract word meanings (Hoffman, 2016; Mestres-Missé et al., 2014). Similarly, children with learning disabilities also learned concrete and abstract words better with pictures via the keyword strategy than with definition rehearsal (Mastropieri et al., 1990). In contrast to the study of Shen (2010) on foreign language learning in adults, children may not automatically imagine pictures for concrete and abstract words as the words were unfamiliar and not linked to a translation, which is why pictures were beneficial for both types of words.

Word concreteness did not influence the group differences in the retention of the knowledge. Taking word concreteness into account, however, revealed better retention in the context group than the picture or control group for both concrete and abstract word meanings. Similarly, as argued for low comprehenders, context sentences may also have benefitted the consolidation or retrieval of newly acquired representations from the mental lexicon or could be used as a mnemonic. Although concrete words were overall learned and retained better than abstract words in our study, no differences were observed in the contextual support effect. Following the context availability theory (Schwanenflugel & Stowe, 1989), additional contextual information might have benefitted retrieval because the representations of both types of words were still impoverished after the short instructional period. In terms of the mnemonic function of context sentences, it could be speculated that context sentences may require the inherent

activation of pictures for both concrete and abstract words, whereas real pictures only supported direct learning of abstract words. By this, the effort during learning may not differ between concrete and abstract words and leads to better retention for both (Bjork & Kroll, 2015).

4.4 | Limitations and future research suggestions

The first limitation of our study is the low teaching intensity, which might have covered group differences. The effects were small, possibly as a result of the low learning gains and need to be confirmed by future investigations. Second, we did not control the time each child spent on studying the words. Therefore, it is unclear if each child actively perceived all the provided information. Finally, we did not explicitly raise the attention of the participants to the link between the meaning and the word forms, which is, however, relevant for successful word learning (Perfetti, 2017; Perfetti et al., 2005). In these respects, the similarity between the word forms of some items could have contributed to confusions. Future investigations should take those above-mentioned aspects into account and other word and child characteristics such as differences between nouns, verbs, and adjectives or differences between learning styles (e.g., Blazhenkova, Becker, & Kozhevnikov, 2011; Elleman et al., 2017). Although this study explored the moderational role of reading comprehension, future research should consider the specific group of children with low reading comprehension abilities but intact decoding skills (poor comprehenders) as well as the heterogeneity within this group to develop efficient vocabulary instructions (e.g., Colenbrander, Kohlen, Smith-Lock, & Nickels, 2016). More repetitions, larger learner involvement, and a smaller word set could intensify the experimental conditions, which could lead to larger effects (Beck & McKeown, 2007; Coyne et al., 2009; Laufer & Rozovski-Roitblat, 2015). To facilitate contextual support effects, the learner involvement and effort could be increased by asking participants to infer the meaning from context precedent to the provision of the definition (Bjork & Kroll, 2015; Golonka et al., 2015; Laufer & Rozovski-Roitblat, 2015). This could be further supported by offering various context sentences (Bolger et al., 2008). To differentiate between overnight learning and long-term retention, future studies should include longer follow-up periods. With respect to the depth of the acquired semantic representations, transfer to reading comprehension should also be considered (e.g., Cain & Oakhill, 2014). It would be interesting to further investigate contextual support effects when children are instructed to imagine pictures resembling the context sentences. This ability might differ between children with various reading comprehension abilities as well as between concrete and abstract words.

5 | CONCLUSION AND IMPLICATIONS

In contrast to context sentences, adding pictures to definitions enhanced direct word learning of children with various reading comprehension abilities on short term better than adding context sentences. Although this effect disappeared over time, pictures

fostered the retention of the acquired knowledge for low comprehenders over and above definitions. Context sentences were not supportive during direct word learning but seemed to help all learners to at least retain what they have acquired during the experiment, particularly children with low reading comprehension abilities. Although a pictorial support effect over and above definitions was only found for low comprehenders in the retention of the knowledge, the fact that pictures supported direct word learning better than context sentences stresses the usefulness to provide pictures during explicit teaching of new vocabulary in a first language learning setting and in classrooms with large individual differences between children. This highlights also the broad utility of multimedia instruction at school. Our study results should encourage teachers to provide rich semantic information during explicit vocabulary instruction at school, which is especially relevant for children with low reading comprehension abilities due to their problems in learning words naturally from context.

FUNDING INFORMATION

This work was supported by the Netherlands Initiative for Education Research (National Regieorgaan Onderwijsonderzoek) as unit of the Netherlands Organisation for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) under Grant NWO 405-15-540-079.

DATA ACCESSIBILITY STATEMENT

Data available on request from the authors.

ORCID

Sophie Gruhn  <https://orcid.org/0000-0003-1845-8717>

Eliane Segers  <https://orcid.org/0000-0002-8616-6890>

Ludo Verhoeven  <https://orcid.org/0000-0002-5108-4705>

REFERENCES

- Acha, J. (2009). The effectiveness of multimedia programmes in children's vocabulary learning. *British Journal of Educational Technology*, 40, 23–31. <https://doi.org/10.1111/j.1467-8535.2007.00800.x>
- Baddeley, A. (1999). *Human memory* (2nd ed.). Boston, MA: Allyn & Bacon.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158–173. <https://doi.org/10.1037/0033-295X.105.1.158>
- Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*, 107, 251–271. <https://doi.org/10.1086/511706>
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). New York, NY: Guilford Press.
- Biemiller, A. (2006). Vocabulary development and instruction: A prerequisite for school learning. In D. K. Dickinson, & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2) (pp. 41–51). New York, NY: The Guilford Press.
- Bjork, R. A., & Kroll, J. F. (2015). Desirable difficulties in vocabulary learning. *The American Journal of Psychology*, 128, 241–252. <https://doi.org/10.5406/amerjpsyc.128.2.0241>
- Blazhenkova, O., Becker, M., & Kozhevnikov, M. (2011). Object-spatial imagery and verbal cognitive styles in children and adolescents: Developmental trajectories in relation to ability. *Learning and Individual Differences*, 21, 281–287. <https://doi.org/10.1016/j.lindif.2010.11.012>
- Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes*, 45, 122–159. <https://doi.org/10.1080/01638530701792826>
- Brus, B. T., & Voeten, M. J. (1979). *Eén-Minuu-Test, vorm A en B. Verantwoording en handleiding* (2nd ed.) [One-Minute-Test, form A and B. Justification and manual]. Nijmegen, The Netherlands: Berkhout Testmateriaal. Lisse, The Netherlands: Swets & Zeitlinger.
- Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica*, 150, 80–84. <https://doi.org/10.1016/j.actpsy.2014.04.010>
- Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods*, 39, 510–526. <https://doi.org/10.3758/BF03193020>
- Bus, A. G., & Neuman, S. B. (2009). *Multimedia and literacy development: Improving achievement of young learners* (1st ed.). New York, NY: Routledge, Taylor & Francis.
- Busch, R. M., Farrell, K., Lisdahl-Medina, K., & Krikorian, R. (2005). Corsi block-tapping task performance as a function of path configuration. *Journal of Clinical and Experimental Neuropsychology*, 27, 127–134. <https://doi.org/10.1080/138033990513681>
- Cain, K., & Oakhill, J. (2011). Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities*, 44, 431–443. <https://doi.org/10.1177/0022219411410042>
- Cain, K., & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Année Psychologique*, 114, 647–662. <https://doi.org/10.4074/S0003503314004035>
- Cain, K., Oakhill, J. V., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language*, 30, 681–694. <https://doi.org/10.1017/S0305000903005713>
- Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, 96, 671–681. <https://doi.org/10.1037/0022-0663.96.4.671>
- Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14, 5–26. <https://doi.org/10.1023/A:1013176309260>
- Carpenter, S. K., & Olson, K. M. (2012). Are pictures good for learning new vocabulary in a foreign language? Only if you think they are not. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 92–101. <https://doi.org/10.1037/a0024828>
- Choi, S. (2016). Effects of L1 and L2 glosses on incidental vocabulary acquisition and lexical representations. *Learning and Individual Differences*, 45, 137–143. <https://doi.org/10.1016/j.lindif.2015.11.018>
- Cito. (2018a, September 14). *Educational measurement*. Retrieved from <https://www.cito.com/we-are/educational-measurement>
- Cito. (2018b, September 14). *LVS-toetsen: Betrouwbaar en diepgaand inzicht in de ontwikkeling van ál je leerlingen* [Pupil- and instruction monitoring test system: Valid and deep insight in the development

- of all your pupils]. Retrieved from <https://www.cito.nl/onderwijs/primair-onderwijs/lvs-toetsen>
- Cohen, M. T., & Johnson, H. L. (2011). Improving the acquisition of novel vocabulary through the use of imagery interventions. *Early Childhood Education Journal*, 38, 357–366. <https://doi.org/10.1007/s10643-010-0408-y>
- Colenbrander, D., Kohnen, S., Smith-Lock, K., & Nickels, L. (2016). Individual differences in the vocabulary skills of children with poor reading comprehension. *Learning and Individual Differences*, 50, 210–220. <https://doi.org/10.1016/j.lindif.2016.07.021>
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407–428. <https://doi.org/10.1037/0033-295X.82.6.407>
- Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, 125, 452–465. <https://doi.org/10.1016/j.cognition.2012.07.010>
- Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R. Jr., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth versus depth. *The Elementary School Journal*, 110, 1–18. <https://doi.org/10.1086/598840>
- Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society*, 364, 3773–3800. <https://doi.org/10.1098/rstb.2009.0111>
- De Paula, J. J., Malloy-Diniz, L. F., & Romano-Silva, M. A. (2016). Reliability of working memory assessment in neurocognitive disorders: A study of the digit span and Corsi block-tapping tasks. *Revista Brasileira de Psiquiatria*, 38, 262–263. <https://doi.org/10.1590/1516-4446-2015-1879>
- Elleman, A. M., Steacy, L. M., Olinghouse, N. G., & Compton, D. L. (2017). Examining child and word characteristics in vocabulary learning of struggling readers. *Scientific Studies of Reading*, 21, 133–145. <https://doi.org/10.1080/10888438.2016.1265970>
- Elliott, R. T., & Zhang, Q. (1998). Interference in learning context-dependent words. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 18, 5–25. <https://doi.org/10.1080/0144341980180101>
- Feenstra, H., Kleintjes, F., Kamphuis, F., & Krom, R. (2010). *Leerling- en onderwijsvolgsysteem. Begrijpend lezen. Groep 3 t/m 6* [Pupil- and instruction monitoring system. Reading comprehension. Grade 1 to 4]. Arnhem, The Netherlands: Cito groep.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London, England: Sage Publications.
- Fukink, R. G. (2005). Deriving word meaning from written context: A process analysis. *Learning and Instruction*, 15, 23–43. <https://doi.org/10.1016/j.learninstruc.2004.12.002>
- Golonka, E., Bowles, A., Silbert, N., Kramasz, D., Blake, C., & Buckwalter, T. (2015). The role of context and cognitive effort in vocabulary learning: A study of intermediate-level learners of Arabic. *The Modern Language Journal*, 99, 19–39. <https://doi.org/10.1111/modl.12191>
- Hoffman, P. (2016). The meaning of 'life' and other abstract words: Insights from neuropsychology. *Journal of Neuropsychology*, 10, 317–343. <https://doi.org/10.1111/jnp.12065>
- Hoffman, P., Jefferies, E., & Lambon Ralph, M. A. (2010). Ventrolateral prefrontal cortex plays an executive regulation role in comprehension of abstract words: Convergent neuropsychological and rTMS evidence. *Journal of Neuroscience*, 30, 15450–15456. <https://doi.org/10.1523/JNEUROSCI.3783-10.2010>
- Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed.). *Cambridge handbooks in psychology* (pp. 247–262). Cambridge, England: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369>
- Kessels, P., Van Zandvoort, M., Postma, A., Kapelle, L., & De Haan, H. (2000). The Corsi block-tapping task: Standardization and normative data. *Applied Neuropsychology*, 7, 252–258. https://doi.org/10.1207/S15324826AN0704_8
- Kim, D., & Gilman, D. A. (2008). Effects of text, audio, and graphic aids in multimedia instruction for vocabulary learning. *Educational Technology & Society*, 11, 114–126. Retrieved from https://www.j-ets.net/ETS/journals/11_3/1.pdf
- Kolich, E. M. (1991). Effects of computer-assisted vocabulary training on word knowledge. *Journal of Educational Research*, 84, 177–182. <https://doi.org/10.1080/00220671.1991.10886012>
- Kort, W., Schittekatte, M., Dekker, P. H., Verhaeghe, P., Compaan, E. L., Bosmans, M., & Vermeir, G. (2005). WISC-III NL Wechsler Intelligence Scale for Children. David Wechsler. *Derde Editie NL. Handleiding en Verantwoording*. [3rd Dutch ed., Manual and justification]. Amsterdam, The Netherlands: Harcourt Test Publishers, NIP Dienstencentrum.
- Laufer, B., & Rozovski-Roitblat, B. (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. *Language Teaching Research*, 19, 687–711. <https://doi.org/10.1177/1362168814559797>
- Long, J. A. (2018). *jtools: Analysis and presentation of social scientific data_*. R package version 1.1.0. Retrieved from <https://cran.r-project.org/package=jtools>
- Mann, B., Newhouse, P., Pagram, J., Campbell, A., & Schulz, H. (2002). A comparison of temporal speech & text cueing in educational multimedia. *Journal of Computer Assisted Learning*, 18, 296–308. <https://doi.org/10.1046/j.0266-4909.2002.00241.x>
- Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning. *Review of Educational Research*, 80, 300–335. <https://doi.org/10.3102/0034654310377087>
- Mastropieri, M. A., Scruggs, T. E., & Mushinski Fulk, B. J. (1990). Teaching abstract vocabulary with the keyword method: Effects on recall and comprehension. *Journal of Learning Disabilities*, 23, 92–96. <https://doi.org/10.1177/002221949002300203>
- Mayer, R. E. (Ed.) (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). In *Cambridge handbooks in psychology*. Cambridge, England: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369>
- McKenzie, S., & Eichenbaum, H. (2011). Consolidation and reconsolidation: Two lives of memories? *Neuron*, 71, 224–233. <https://doi.org/10.1016/j.neuron.2011.06.037>
- McKeown, M. G., & Beck, I. L. (2004). Direct and rich vocabulary instruction. In J. F. Baumann, & E. J. Kame'enui (Eds.), *Vocabulary instruction* (pp. 13–27). New York, NY: Guilford Press.
- Meisel, J. (2011). *First and second language acquisition: Parallels and differences* (1st ed.). Cambridge, England: Cambridge University Press. <https://doi.org/10.1017/CBO9780511862694>
- Mestres-Missé, A., Münte, T. F., & Rodríguez-Fornells, A. (2014). Mapping concrete and abstract meanings to new words using verbal contexts. *Second Language Research*, 30, 191–223. <https://doi.org/10.1177/0267658313512668>
- Mondria, J.-A., & Wit-de Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics*, 12, 249–267. <https://doi.org/10.1093/applin/12.3.249>

- Nash, H., & Snowling, M. (2006). Research report: Teaching new words to children with poor existing vocabulary knowledge: A controlled evaluation of the definition and context methods. *International Journal of Language & Communication Disorders*, 41, 335–354. <https://doi.org/10.1080/13682820600602295>
- Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition*, 70, B1–B13. [https://doi.org/10.1016/S0010-0277\(99\)00004-9](https://doi.org/10.1016/S0010-0277(99)00004-9)
- National Reading Panel. (2000). Comprehension—Part 1: Vocabulary instruction. In Author (Ed.), *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (pp. 15–35). Retrieved from <https://www1.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf>
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45, 255–287. <https://doi.org/10.1037/h0084295>
- Pallant, J. (2013). Checking the reliability of a scale. In Author (Ed.), *SPSS survival manual* (5th ed.) (pp. 95–99). Maidenhead, England: Open University Press.
- Perfetti, C. A. (2007). Reading ability: From lexical quality to comprehension. *Scientific Studies of Reading*, 11, 357–383. <https://doi.org/10.1080/10888430701530730>
- Perfetti, C. A. (2017). Lexical quality revisited. In E. Segers, & P. van den Broek (Eds.), *Developmental perspectives in written language and literacy* (1st ed.) (pp. 51–68). Amsterdam, The Netherlands: John Benjamins. <https://doi.org/10.1075/z.206>
- Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1281–1292. <https://doi.org/10.1037/0278-7393.31.6.1281>
- Pittelman, S. D., Heimlich, J. E., Berglund, R. L., & French, M. P. (1991). *Semantic feature analysis: Classroom applications*. Newark, DE: International Reading Association.
- Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior*, 19, 221–243. [https://doi.org/10.1016/S0747-5632\(02\)00015-8](https://doi.org/10.1016/S0747-5632(02)00015-8)
- Read, J. (2004). Plumbing the depth: How should the construct of vocabulary knowledge be defined? In P. Bogaards, & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (1st ed.). *Language Learning & Language Teaching* (pp. 209–227). Amsterdam, The Netherlands: John Benjamins.
- Sadoski, M. (2005). A dual coding view of vocabulary learning. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 21, 221–238. <https://doi.org/10.1080/10573560590949359>
- Schwanenflugel, P. J., & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly*, 24, 114–126. <https://doi.org/10.3758/BF03208259>
- Shen, H. H. (2010). Imagery and verbal coding approaches in Chinese vocabulary instruction. *Language Teaching Research*, 14, 485–499. <https://doi.org/10.1177/1362168810375370>
- Silverman, R., & Hines, S. (2009). The effects of multimedia-enhanced instruction on the vocabulary of English-language learners and non-English-language learners in pre-kindergarten through second grade. *Journal of Educational Psychology*, 101, 305–314. <https://doi.org/10.1037/a0014217>
- Smith, B. D., Stahl, N., & Neil, J. (1987). The effect of imagery instruction on vocabulary development. *Technical Report*, 87, 1–23. <https://doi.org/10.1080/10790195.1987.10849909>
- Stafura, J. Z., & Perfetti, C. A. (2017). Integrating word processing with text comprehension: Theoretical frameworks and empirical examples. In K. Cain, D. L. Compton, & R. K. Parrilla (Eds.), *Theories of reading development* (1st ed., Vol. 15). *Studies in Written Language and Literacy* (pp. 9–31). Amsterdam, The Netherlands: John Benjamins.
- Stahl, S. (1983). Differential word knowledge and reading comprehension. *Journal of Reading Behavior*, 15, 33–50. <https://doi.org/10.1080/10862968309547495>
- Stahl, S. A., & Fairbanks, M. M. (2006). The effects of vocabulary instruction: A model-based meta-analysis. In K. A. D. Stahl, & M. C. McKenna (Eds.), *Reading research at work: Foundations of effective practice* (1st ed.) (pp. 226–261). New York, NY: Guilford.
- Staphorsius, G., & Krom, R. (1998). *Toetsen Begrijpend Lezen. Handleiding* (3rd ed.). [Reading comprehension test. Manual]. Arnhem, The Netherlands: Cito groep.
- Staphorsius, G., Krom, R., Kleintjes, F., & Verhelst, N. (2004). *Toetsen begrijpend lezen. Verantwoording. Verslag van het kalibratie-, validerings- en normeringsonderzoek* [Reading comprehension tests. Scientific justification. Research report on calibration, validation, and standardization]. Arnhem, The Netherlands: Citogroep.
- Stichting Digiwak. (2017, February 10). *DigiWak*. Retrieved from <https://www.digiwak.nl/>
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 295–312. [https://doi.org/10.1016/0959-4752\(94\)90003-5](https://doi.org/10.1016/0959-4752(94)90003-5)
- Tabachnik, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.
- Tellings, A., Hulsbosch, M., Vermeer, A., & Bosch, A. V. (2014). BasiLex: An 11.5 million words corpus of Dutch texts written for children. *Computational Linguistics in the Netherlands Journal*, 4, 191–208. Retrieved from <http://repository.uibn.ru.nl/bitstream/handle/2066/134845/134845.pdf?sequence=1>
- Tomesen, M., Weekers, A., Hiddink, L., & Jolink, A. (2017). *Wetenschappelijke verantwoording. Begrijpend lezen 3.0 voor groep 6* [Scientific justification. Reading comprehension test version 3.0 for grade 4]. Arnhem, The Netherlands: Cito. Retrieved from <https://www.cito.nl/-/media/Files/kennisbank/cito-bv/17-cito-wv-begrijpend-lezen-3-0-groep-6.pdf?la>
- Van Dale. (2017, February 7). *Dikke van Dale*. Retrieved from <http://surfdiensten2.vandale.nl/zoeken/zoeken.do>
- van der Ven, F., Takashima, A., Segers, E., & Verhoeven, L. (2017). Semantic priming in Dutch children: Word meaning integration and study modality effects. *Language Learning: A Journal of Research in Language Studies*, 67, 546–568. <https://doi.org/10.1111/lang.12235>
- Verhoeven, L., Van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, 15, 8–25. <https://doi.org/10.1080/10888438.2011.536125>
- Verhoeven, L., & Vermeer, A. (1993). *Taaltoets Allochtone Kinderen: Diagnostische toets voor de vaardigheid Nederlands bij allochtone en autochtone kinderen in de bovenbouw van het basisonderwijs* [Language Test for Allochthone Children: Diagnostical Dutch ability test for allochthone and autochthone children in final primary school grades]. Tilburg, The Netherlands: Zwijsen uitgeverij.
- Verhoeven, L., & Vermeer, A. (1996). *Taalvaardigheid in de bovenbouw: Nederlands van autochtone en allochtone leerlingen in het basis- en mlk-onderwijs*. *Studies in meertaligheid 7*. [Language ability in the upper primary grades: Dutch from autochthonous and allochthonous pupils in the mainstream and special needs primary education]. Tilburg, The Netherlands: University Press.
- Verhoeven, L., & Vermeer, A. (1999). *Handleiding Leeswoordenschat* [Manual for the test Reading Vocabulary]. Arnhem, The Netherlands: Cito groep.

- Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in Psychology, 4*(6), 1–8. <https://doi.org/10.3389/fpsyg.2012.00046>
- Webb, S. (2007a). Learning word pairs and glossed sentences: The effects of a single context on vocabulary knowledge. *Language Teaching Research, 11*, 63–81. <https://doi.org/10.1177/1362168806072463>
- Webb, S. (2007b). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28*, 46–65. <https://doi.org/10.1093/applin/aml048>
- Webb, S., & Nation, I. S. P. (2017). Vocabulary size and growth. In Authors (Ed.), *How vocabulary is learned* (3rd chapter ed.). Oxford: Oxford University Press.
- Winkler Prins. (2012). *Jeugdwoordenboek Nederlands* (2nd ed.) [Youth lexicon Dutch]. Houten - Antwerpen, Belgium: Uitgeverij Unieboek. Het Spectrum.
- Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis. *Computer Assisted Language Learning, 24*, 39–58. <https://doi.org/10.1080/09588221.2010.523285>
- Zhou, N., & Yadav, A. (2017). Effects of multimedia story reading and questioning on preschoolers' vocabulary learning, story comprehension and reading engagement. *Educational Technology Research and Development, 65*, 1523–1545. <https://doi.org/10.1007/s11423-017-9533-2>

How to cite this article: Gruhn S, Segers E, Verhoeven L. Moderating role of reading comprehension in children's word learning with context versus pictures. *J Comput Assist Learn.* 2020;36:29–45. <https://doi.org/10.1111/jcal.12387>

APPENDIX A.

OVERVIEW OF THE STUDY MATERIAL

TABLE A1 Material for the experiment: Concrete items

Item	Definition	Context sentence	Picture description
Vesting <i>Fortress</i>	Plaats die door dikke muren is omringd <i>Place which is surrounded by thick walls</i>	De ridder verdedigt de vesting van de stad. <i>The knight defends the fortress of the city.</i>	Fortress on a hill
Scheg <i>Bread</i>	Dikke stuk of dikke snee brood <i>Thick piece or thick slice of bread</i>	Hij snijdt een scheg van het brood af. <i>He cuts off a thick slice of the bread.</i>	Several thick slices of bread
Girandole <i>Girandole</i>	Kandelaar met meerdere armen <i>Candleholder with several branches</i>	Op de tafel branden de kaarsen in de girandole. <i>On the table, the candles are burning in the girandole.</i>	Girandole with lighted candles
Keen <i>Fissure</i>	Barst of scheur in de huid <i>Rip or fissure on the skin</i>	Door de kou kreeg hij een keen in zijn hand. <i>From the cold, he got a fissure on his hand.</i>	Heel with small fissures
Talisman <i>Talisman</i>	Object dat geluk brengt en onheil afweert <i>Object that brings fortune and protects against harm</i>	Met de talisman in haar hand kreeg ze vertrouwen. <i>With the talisman in her hand, she got confidence.</i>	Little dolls often used as talisman in the Netherlands
Spant <i>Timber</i>	Balken van een dak- of scheepsgeraamte <i>Bar of the timberwork from a roof or ship</i>	Het oude schip had houten spanten en een zeil. <i>The old ship had wooden timber and a rope.</i>	Timberwork of a boat
Majuskel <i>Majuscule</i>	Naam van hoofdletters in oude handschriften <i>Name of the capital letters in old scripts</i>	In het oude boek waren vooral de majuskels versierd. <i>In the old book, mainly the majuscules were ornate.</i>	Page with old script and stressed capital letters
Strapontijn <i>Folding seat in a vehicle</i>	Neerklapbare of vouwbare zitting in een rijtuig <i>Seat in a vehicle that can be folded or collapsed</i>	In de bus was alleen een plek op een strapontijn vrij. <i>In the bus, was only a place on a folding seat available.</i>	Folding seats in a bus
Contour <i>Contour</i>	Geschetste vorm of omtreklijn van iets <i>Approximated form or outline shape of something</i>	Hij ziet de contour van zijn lichaam in de schaduw. <i>He sees the contour of his body in the shadow.</i>	Contour line of a dancer
Flink <i>Flink</i>	Zijkant van een voorwerp of levend wezen <i>Side of an object or creature</i>	Hij drukte zijn hak in de flank van het paard. <i>He pressed his heels in the flank of the horse.</i>	Horse from side perspective

Note. The English translation is provided in italics.

TABLE A2 Material for the experiment: Abstract words

Item	Definition	Context sentence	Picture description
Decennium <i>Decennium</i>	Tijdafstand of tijdsverschil van precies tien jaar <i>10 years' time difference or distance</i>	Zij houdt van de muziek uit het laatste decennium. <i>She loves the music of the last decennium.</i>	Time line from year 2000 to 2010 to 2020
Compensatie <i>Compensation</i>	Iets wat iemand als vergoeding terug krijgt <i>Something that you receive for damages</i>	Als compensatie voor de vertraging kreeg hij zijn geld terug. <i>As compensation for the delay of the train, he received money back</i>	Hand with money
Schikking <i>Accord with compromises</i>	Overeenkomst waarbij van beide partijen wat wordt toegegeven <i>Agreement between two parties where both admit something</i>	De ruziemakende burens kwamen tot een schikking. <i>The contesting neighbors accorded with compromises.</i>	Two people are sitting opposite of each other at a table in a business meeting to discuss. Their hands are placed on the table with each crossed fingers.
Prestige <i>Prestige</i>	Respect krijgen door goede eigenschappen en prestaties <i>Receive respect by good traits and achievements</i>	Hij wilde niet zijn goede naam en prestige verliezen. <i>He did not want to lose his reputation and prestige.</i>	A conceited looking person in a suit is filmed and interviewed.
Preventie <i>Prevention</i>	Voorzorgsmaatregelen om iets negatiefs te voorkomen <i>Precautionary principles to avoid something negative</i>	Ze vonden een middel voor de preventie van longontsteking. <i>They found a medicine for the prevention of lung infection.</i>	Someone received a vaccination
Charisma <i>Charisma</i>	Positieve uitstraling van mensen op anderen <i>Positive vibes of people and their effect on others</i>	De spreker kreeg veel aandacht door zijn charisma. <i>The speaker got a lot of attention due to his charisma.</i>	Barack Obama
Privilege <i>Privilege</i>	Iets wat iemand wel mag, maar anderen niet <i>Something that someone is permitted to do, but others not</i>	De actrice heeft het privilege om met chauffeur te reizen. <i>The actress has the privilege to travel with a chauffeur.</i>	Sign used in busses to provide seats to persons with problems to stand and walk
Obsessie <i>Obsession</i>	Bijna nergens anders meer aan kunnen denken <i>Almost not being able to think about something else</i>	John had een obsessie voor auto's gekregen. <i>John got an obsession with cars.</i>	Group of soccer fans
Verontwaardiging <i>Feelings of anger and consternation</i>	Gevoel van kwaadheid door oneerlijke behandeling <i>Angry feeling about being treated unfair</i>	Toen ze dat hoorde, stakte haar adem van verontwaardiging. <i>When she heard about this, her breath caught of anger and consternation.</i>	Two youngsters look at a smart phone screen. Their facial expressions and gestures show feelings of anger and consternation about what they have read.
Duplicaat <i>Duplicate</i>	Kopie of verdubbeling van een origineel <i>Copy or doubling of an original</i>	De kinderen kregen een duplicaat van de huissleutel. <i>The children got a duplicate of the flat key.</i>	Two identical keys

Note. The English translation is provided in italics.

APPENDIX B.

PARTICIPANT INFORMATION AND COMPARISON OF GROUPS (CONTEXT, PICTURE, AND CONTROL)

TABLE B1 Mean scores for each measure and scores of the group comparison

Measure and test	Context group	Picture group	Control group	Test on group differences	
				Test statistic	p
<i>n</i>	63	66	60	—	—
Age in months	118.86 (5.71)	118.53 (5.05)	117.45 (6.00)	$\chi^2(2) = 1.87$.39
Visual memory					
Corsi block-tapping task [72]	46.70 (13.50)	43.45 (15.42)	48.65 (18.96)	$\chi^2(2) = 1.63$.44
Verbal memory					
WISC digit span forwards [16]	7.65 (1.99)	8.00 (1.61)	7.65 (1.56)	$\chi^2(2) = 2.14$.34
WISC digit span backwards [14]	4.98 (1.66)	4.82 (1.50)	4.65 (1.59)	$\chi^2(2) = 1.36$.51
Reading fluency					
One-minute test [116]	60.97 (15.80)	61.56 (13.08)	63.62 (12.96)	$F(2, 186) = 0.60$.55
Vocabulary					
Nationally standardized test [150] ^a	29.45 (7.36)	31.13 (6.00)	29.00 (7.70)	$F(2, 177) = 1.16$.32
Reading vocabulary [50] ^b	74.39 (14.67)	79.33 (13.85)	78.08 (16.85)	$\chi^2(2) = 1.01$.61
Passive vocabulary breadth [24,000] ^c	6241.27 (3638.00)	7178.13 (3529.49)	6728.81 (3545.71)	$F(2, 183) = 1.09$.34
Reading comprehension					
Nationally standardized test [369] ^d	169.66 (31.40)	172.96 (20.03)	169.28 (33.97)	$\chi^2(2) = 0.56$.76

Note. Standard deviations are presented within parentheses. Maximum test scores are presented within square brackets. The group differences were tested with one-way ANOVAs (F). If the assumptions of one-way ANOVAs were not met, we used Kruskal–Wallis tests (χ^2 ; Field, 2009).

^aTest from Verhoeven and Vermeer (1999) as implemented in the Dutch schools' pupil and instruction monitoring system (Cito, 2018a, 2018b).

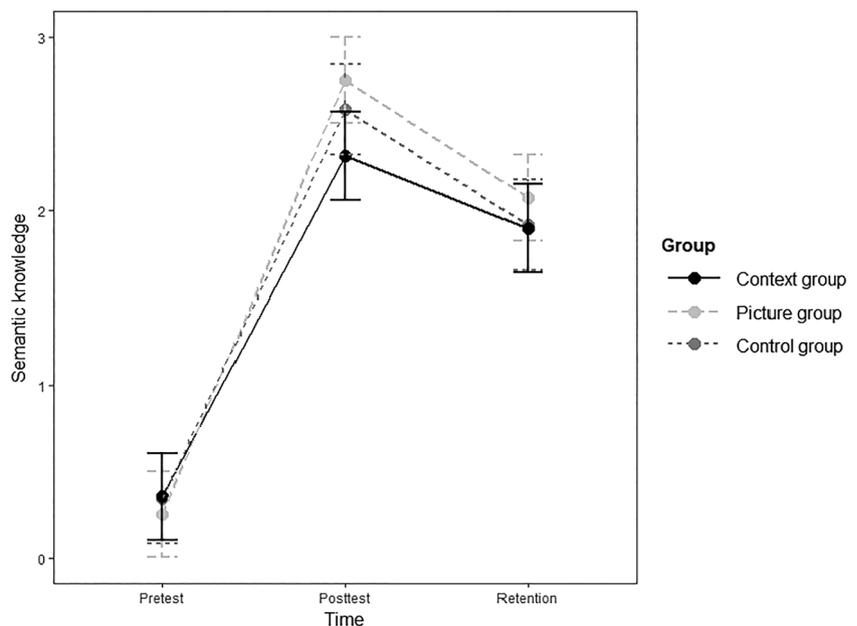
^bDutch standardized test from Verhoeven and Vermeer (1993), which is similar to the test described in a.

^cPredicted from vocabulary tests a and b, see test manuals for further details.

^dTest editions from Feenstra et al. (2010) and Staphorsius and Krom (1998) as implemented in the Dutch schools' pupil and instruction monitoring system (Cito, 2018a, 2018b).

APPENDIX C.

ADDITIONAL FIGURES

**FIGURE C1** Mean sum scores of the semantic knowledge (square root transformed) for each group over time