DO DICTIONARY DEFINITIONS SUPPORT SEVENTH GRADERS' ACQUISITION OF WORD MEANING WHILE READING?

JULIA HODSON AND URSINA MARKWALDER

Fachdidaktische Forschung, Pädagogische Hochschule Zürich

Abstract

Looking up unknown words in a dictionary while reading is a common recommendation at school. However, little is known about the effects of dictionary use on the acquisition of word meaning while reading in the language of instruction. This study investigated whether students acquire the meaning of unknown words by using a dictionary while reading. The aim was also to explore whether students´ knowledge of the target words differ regarding the dictionary used (print or electronic dictionary). We conducted an experiment in which seventh-grade students in Germany (*N* = 153) were asked to read a text in the language of instruction (German) and look up twelve target words in a dictionary. Subjects were assigned to three treatment conditions (Kindle/integrated electronic dictionary, Duden/print dictionary, control/no dictionary). Knowledge of the target words was tested by means of a vocabulary test twice (one after reading the text and one 14 days later). Best fitting models were calculated using generalized linear mixed models (GLMM). For the first vocabulary test, results showed a significant effect of treatment condition. However, contrasts revealed a significant difference only between the Kindle group and the Duden group and between the Kindle group and the control group. No differences between groups could be found for the second vocabulary test. The results support the use of dictionary definitions for the acquisition of word meanings while reading on electronic reading devices.

Keywords: dictionary use, word meaning acquisition, vocabulary learning, electronic reading devices

Hodson, J., & Markwalder, U. (2025). Do dictionary definitions support seventh graders' acquisition of word meaning while reading? L1-Educational Studies in Language and Literature, 25, 1-18. https://doi.org/10.21248/l1esll.2025.25.1.744

Corresponding author: Julia Hodson, Fachdidaktische Forschung, Pädagogische Hochschule Zürich, email: Julia.hodson@phzh.ch

1. INTRODUCTION

Dictionaries are commonly considered as an important source of information about word meanings (Stahl & Nagy, 2006, p. 182). When coming across an unknown word while reading, using a dictionary is therefore a frequently used word-learning strategy not only in the context of vocabulary learning in a foreign language but also in the language of instruction (Graves et al., 2018, p. 533). Consequently, it is a common practice for teachers to encourage students to look up unfamiliar words while reading. Using dictionary definitions to look up unfamiliar words while reading is likewise a strategy that students learn in reading strategy training programs to prevent or solve reading comprehension problems. Although dictionaries are considered as useful tools, little is known about the effects of dictionary use while reading on vocabulary acquisition in the language of instruction.

1.1 Vocabulary learning through reading

Previous research has shown that readers acquire vocabulary during reading either incidentally (Nagy et al., 1985, 1987; Herman et al., 1987; Schwanenflugel et al., 1997; Swanborn & de Glopper, 1999) or intentionally (Carnine et al., 1984; Jenkins et al., 1984; McKeown, 1985; Fukkink, 2005) by deriving word meanings from context (Swanborn & de Glopper, 1999, p. 275). Prior research has also revealed an interaction between reading comprehension and vocabulary acquisition during reading: Good readers learn more words while reading than poor readers (Cain et al., 2004; Herman et al., 1987; Jenkins et al., 1989; Swanborn & de Glopper, 2002; Swanborn & de Glopper, 1999). Nagy et al. (1987) identified further variables that influence the process of word learning during reading, including the readability of the text, the density of unknown words, and the difficulty in conceptualizing the meaning of a word (Nagy et al., 1987). There also seems to be an influence of age, stating that older students learn more new word meanings while reading than younger students (Carnine et al., 1984; Fukkink et al., 2001; Merisuo-Storm & Soininen, 2010). Whether these differences in age result from reading skills improving over the years, general cognitive development, growing vocabulary, or a combination of these factors is still unclear.

1.2 Dictionary definitions and vocabulary learning

Studies conducted by Bolger et al. (2008) and Nist and Olejnik (1995) compared the impact of dictionary definitions on word learning to incidental word learning. Using definitions in addition to contexts in which unknown words were used, college students showed significantly better results in tests of word knowledge (Bolger et al., 2008; Nist & Olejnik, 1995). Studies investigating effects of using definitions while

reading on learning new words in a foreign language have also shown significantly better results when using a dictionary compared to incidental word learning (Alharbi, 2016; Flynn, 2007; Liu & Lin, 2011; McCreary & Dolezal, 1999; Knight, 1994; Luppescu & Day, 1993). However, no differences were found in a study by Fischer (1994), who argues that "It may be difficult for students to map a definition onto contextual information and to integrate both types of information into a coherent meaning representation" (1994, p. 571).

Research on the usefulness of dictionary definitions for word learning irrespective of reading has shown that young children experience great difficulties in understanding definitions adequately (McKeown 1985; 1993; Miller & Gildea, 1985; Scott and Nagy,1997). These studies established a critical perspective towards the use of dictionary definitions at schools for vocabulary acquisition. Stahl and Nagy (2006) named three limitations of dictionary definitions: dependency of a words meaning on context, lack of information about a word's usage, and difficulties children show in understanding dictionary definitions (p. 182). Whereas the authors (Stahl & Nagy, 2006) "don´t see these limitations as reasons to give up the use of dictionaries" (p. 183), Beck et al. (2002) advised against using dictionary definitions for word learning: "The reality is that definitions are not an effective vehicle for learning word meanings" (p. 33).

While this may be true for an isolated use of dictionary definitions for word learning, it is reasonable to argue that the general limitations of dictionary definitions mentioned by Stahl & Nagy (2006), i. e. dependency of a words meaning on context and lack of information about a word's usage, are not in effect when using dictionary definitions while reading. When students look up an unknown word while reading, the unknown word is embedded in a natural context, which means, that students do have context to determine the contextual meaning of the word. Also, students experience an example of the word's usage by reading a natural text in which the word is used. Therefore, it could be hypothesized that dictionary definitions can foster vocabulary acquisition through looking up unknown words while reading.

1.3 Use of electronic dictionaries while reading

In the course of digitization, reading and dictionary use have changed. When reading on electronic reading devices, readers can look up words directly by tapping on the touchscreen. Research has shown that users of electronic dictionaries look up more words (Alharbi, 2016; Aust et al., 1993; Diehr et al., 2013; Koyama & Takeuchi, 2007, 2004a; Liu & Lin, 2011; Reinking & Rickman, 1990) and take less time (Alharbi, 2016; Chen, 2010; Liu & Lin, 2011; Shizuka, 2003; Weschler & Pitts, 2000) than users of print dictionaries. In contrast, no differences in lookup speed were found in studies by Koyama and Takeuchi (2004b) and Nesi (2000) for lookup speed. Also, Nesi (2000) detected no differences in the numbers of lookups.

It can be assumed that the use of electronic dictionaries has a stronger effect on vocabulary acquisition while reading than the use of print dictionaries because of

simple and quick lookup. This hypothesis is based on the theory of the capacity of the working memory (Baddeley, 2012): When using a print dictionary to derive the meaning of an unknown word while reading, the reader must process multiple competing pieces of information in working memory that are not related to word meaning (e.g., What is the lemma of the unknown word? Where do I find the lemma? What is the alphabetic order?). While reading on an electronic reading device and using the integrated dictionary, none of the information related to the process of looking up a word in a print dictionary must be processed. Instead, the reader can concentrate on deriving the meaning of the unknown word by using the dictionary definition and the context. In contrast, the involvement load hypothesis (Hulstijn & Laufer, 2001; Laufer & Hulstijn, 2001; Laufer, 2000) states that use of print dictionaries and related involvement is beneficial for word learning processes because they are more difficult to consult than digital dictionaries (Lew et al., 2014, p. 347). However, studies comparing the use of print dictionaries to electronic ones during reading have shown differing results: In a study conducted by Reinking and Rickman (1990), 11-to-12-year-old participants using an electronic dictionary while reading on their computers performed better in the vocabulary test compared to students using a print dictionary. In addition, participants reading in a foreign language learned more words when looking up words in electronic versus print dictionaries (Alharbi, 2016; Liu & Lin, 2011; Flynn, 2007; Ronald & Tajino, 2005). No differences regarding the effects of electronic and print dictionaries on vocabulary learning in a foreign language were found in studies by Kobayashi (2007), Koyama and Takeuchi (2003) and Osaki et al. (2003). Studies comparing the use of print dictionaries to electronic ones for vocabulary learning in a foreign language beyond reading, have also shown differing results: While studies of Dziemianko (2010, 2017) suggests electronic dictionaries to be more effective regarding long-term learning effects than print dictionaries, these results could not be replicated in a study by Chen (2011).

1.4 Purpose of present study

The present study investigated the effects of dictionary use while reading on the acquisition of word meanings. An extensive literature review (Hodson, 2021) identified only one study (Reinking & Rickman, 1990) that examined the effects of dictionary use on students' word knowledge while reading in the language of instruction. In addition, the present study explored whether using different dictionaries (print dictionary vs. integrated dictionary in electronic reading devices) has an impact on the acquisition of word meanings while reading. Because of studies showing stronger effects of electronic dictionaries on long-term learning effects than print dictionaries (Dziemianko 2010, 2017), it was also tested, whether there is a difference between short- and long-term vocabulary learning effects related to different dictionaries used while reading. Thus, the research questions are as follows:

1) What are the short- and long-term effects of dictionary use while reading on word meaning acquisition?

2) Does using integrated dictionaries while reading on electronic reading devices show stronger short- and long-term effects on the acquisition of word meaning than using print dictionaries?

It was hypothesized that children using a dictionary do better than control group subjects with no dictionary in a vocabulary test on two measurement time points (after the reading-activity + two weeks later). A second hypothesis was that the group reading on an electronic reader and using the integrated dictionary would perform better in a vocabulary test on both measurement time points than the group using a print dictionary. The study aims to empirically validate the recommendation to look up unknown words when reading. Another objective is to explore the potential of digital media for language learning, as it is still unclear whether students' word meanings acquisition during reading benefits from the easy and quick electronic look-up.

2. METHOD

2.1 Design

An experimental research design was chosen in which participants were randomly assigned to the independent variable group. There were three conditions, corresponding to the number of two experimental groups and one control group: Kindle group, Duden group, and control group. The groups differed regarding the medium used for reading and looking up word: Participants assigned to the Kindle group read a text on the electronic *Kindle* reader and looked up words in the digital version of the *Duden Universalwörterbuch* (*DUW*[7]), using the integrated lookup function. Participants assigned to the Duden group read a paper-based copy of the text and consulted the print dictionary *Duden Universalwörterbuch* (*DUW*[7]). The control group read a paper-based copy of the text with no dictionary. After reading, students completed a written vocabulary test on the words looked up. The number of correct solutions in the vocabulary test (mR) is considered the dependent variable. After a time of 14 days had passed, participants took the vocabulary test again to determine long-term learning effects. Since reading comprehension correlates with word learning while reading (Cain et al., 2004; Herman et al., 1987; Jenkins et al., 1989; Swanborn & de Glopper, 2002; Swanborn & de Glopper, 1999), reading comprehension was considered as a control variable. As for many students the language of instruction does not correspond to their family language, i.e. the language they speak with their families at home, family language acted also as a variable to exclude potential differences regarding learning effects related to student’s family language. For this purpose, students were asked to name their family language if it was not German.

2.2 Participants

A total of 153 seventh-grade students from six classrooms across two local comprehensive schools in Germany participated in the study (mean age = 13.24, *SD* = 0.56). These two schools were selected, because they accepted the invitation to participate in the study, which was sent to all comprehensive schools in the area. The sample size of participants was selected a priori based on a power analysis for an *F*-test (ANCOVA) using $p = .05$, *power* = .80, with an expected medium effect size ($f = 0.25$) that recommended an *N* of 118 (G*Power 3; Faul et al., 2007). Grade seven was chosen because the confident use of dictionaries as a working technique should be a prerequisite for this level of schooling. Class membership was not available in the data.

Participation was voluntary; parental consent was obtained in writing. Overall, 111 data sets were included in the inferential statistical analysis. Data sets of subjects were not included if students were absent at any of three measurements (main survey $n = 16$, post-survey $n = 15$, non-participation in the reading comprehension test $n = 8$) or if participants had not completed the vocabulary test ($n = 3$). Although the groups were randomly assigned, a significantly unequal distribution of gender within the groups existed (χ^2 (2) = 6.23, $p = .044$; see Table 1): While girls are overrepresented in the Kindle group, boys are overrepresented in the Duden group, and overrepresented in the control group. Reasons for the unequal gender distribution could not be determined retrospectively. However, the unequal distribution did not result from the reduction of the data sets, since it was already identifiable in the unadjusted data. The distribution of students whose family language is identical to the language of instruction and students whose family language differ from the language of instruction can be neglected (χ^2 (2) = 0.89, p = .641).

Table 1. Participants

	Kindle (n=33)		Duden (n=36)		Control (n=42)		Total (n=111)	
Variable	n	%	n	%	n	%	n	%
Gender								
Girls	22	66.7	14	38.9	18	42.9	54	51.4
Boys	11	33.3	22	61.1	24	57.1	57	48.6
Family language (FL)								
FL = language of instruction	18	54.5	24	66.7	28	66.7	70	63.1
FL ≠ language of instruction	13	39.4	12	33.3	13	30.9	38	34.2
n.a.	2	6.1			1	2.4	3	2.7

2.3 Materials

2.3.1 Reading comprehension

Reading comprehension was measured with the *Reading Speed and Comprehension Test for Grades 6-12* (LGVT 6-12) (Schneider et al., 2007). This test exhibits high reliability, and it was chosen as the sole standardized and readily available assessment

for evaluating seventh-grade reading comprehension during the data collection period. The LGVT 6-12 test manual had a retest reliability of $r = .87$ ($p < .001$; $N = 103$) for reading comprehension (Schneider et al., 2007, p. 17–19). Test validity was indicated by the correlation of $r = .59$ ($N = 711$) with the reading comprehension test of the PISA 2000 field study.

2.3.2 Text and target words

The text selected for the study is a newspaper article on how climate change occurring in the alpine regions affects local tourism (Dambeck, 2008). The original text was simplified for the experiment regarding its length and text structure. This adjustment was made to keep the demands on text comprehension low and allow participants to be able to derive the meaning of unknown words (Nagy et al., 1987, p. 255). To determine words in the text which are unknown to seventh-graders the text was given in a previous evaluation to three teachers and their students ($N = 33$). The teachers were asked to mark words that they expected to be unknown to seventh grade students. The students were asked to read the text and mark words that they don´t know or whose meaning they cannot explain. Subsequently, words that were marked most frequently were selected as target words (5 nouns, 2 verbs, 4 adjectives, 1 idiom). In the experimental text, the target words have been underlined and marked in bold to indicate to participants which words they should look up.

2.3.3 Dictionary

The Kindle group and the Duden group used the German dictionary *Duden Universalwörterbuch* (*DUW*). With its more than 500,000 headwords, *DUW*[7] is a comprehensive and general monolingual dictionary of contemporary German. The *DUW* was selected because the electronic reader *Kindle* offers it as an integrated dictionary. The dictionary entries relevant to this study are identical in content and microstructure in both the printed dictionary and the digital version.

2.3.4 Vocabulary text

Acquisition of word meaning was defined as the ability to identify the meaning of the target word in the vocabulary test as it was used in the reading text. Therefore, the vocabulary test was constructed in a multiple-choice format in which students should choose the meaning of the target word from five different possibilities (one correct answer, four distractors).

For each target word multiple test items were developed. Distractors were lexemes semantically close to the meaning of the target word (i.e., *Entscheidung* [decision] for the target word *Option* [option]), lexemes of the same lexical field (i.e., "Die **massive** Nutzung von Kunstschnee… bedeutet: Die Nutzung von Kunstschnee in begrenztem / kleinem / großem / verschiedenem Umfang" [The **massive** use of artificial

snow... means: the use of artificial snow on a limited / small / large / varying scale.]) or phonologically similar lexemes (i.e., *Optik* [optic] for the target word *Option* [option]). Furthermore, words of the dictionary entry of the target word which could be misunderstood as the meaning of the target word were used as distractors (i.e., *Untersuchung* [investigation] for the target word *Fazit* [conclusion]; dictionary entry: "Das F. der Untersuchungen, Überlegungen war jedes Mal das gleiche" [The conclusion of the investigations, reflections were the same each time] (DUW[7], p. 581)). Content from the study text with no semantic relationship to the target word were also chosen as distractors (i.e., *Klimaveränderung* [climate change] for the target word *Kalkulation* [calculation]).

The test items were reviewed by six experts in the field of language learning and teaching. After the expert-review, one test item for each target word was selected. The revised test was than pretested with seventh graders at two comprehensive schools (N = 99) different from the schools of the main study. In terms of reliability, the pretest achieved an alpha coefficient of α = .63, making the test suitable for this study yet unsuitable for individual diagnosis (Döring & Bortz, 2016, p. 444; Robinson et al., 1991, p. 13). Item difficulty P_i ranged from .27 to .83; three items with low or negative discriminatory power (r_{it} < .25) were revised to increase item discriminatory power and test reliability.

For the main study, two versions of the vocabulary test were created differing in item order. A score of one point was awarded for each correctly solved item, up to a maximum score of 12 points. Incorrect solutions, ungiven answers, and invalid tasks (tasks with more than one check) received zero points.

To study the unidimensionality of the test a confirmatory factor analysis was conducted with two latent factors describing the items at the first and second measurement point, including correlated error terms of the same items over time and using the WLSMV-estimator for dichotomous indicator variables. Results indicated appropriate fit (Hu & Bentler, 1999; χ^2 [197] = 228.66, p = .06; RMSEA = .03, CFI = .96, TLI = .96).

The test achieved an acceptable alpha coefficient of α = .60 for the first measurement point and an alpha coefficient of α = .69 for the second measurement point. Alpha coefficients between .6 to .69 are considered acceptable for tests designed to examine group differences (Robinson et al., 1991, p. 12–13). As mentioned before, low alpha coefficient Item difficulty exhibited an optimal range of spread, ranging from P_i = .29 to .77 for the first measurement point and from $P_{i=}$ .22 to .76 for the second measurement point.

2.4 Procedure

The data collection was conducted in February and March 2016 during two 45-minute lessons. The first lesson was held in the classroom of the participating classes under the guidance of the experimenter. First, the reading comprehension test was administered according to the instructions in the test manual. Afterwards, a small

competition was arranged to review strategies to look up words in printed dictionaries. Participants were instructed to look up as many words as they could from a pool of ten given words in seven minutes. Students than discussed tips and tricks for quick lookups such as using the alphabetical order, head words and markers in the margin. Subsequently, students received a text which contained the ten words looked-up before and were asked to use the dictionary to derive the meaning of the target words in context. This was done to practice using dictionary definitions to derive the meaning of unknown words when reading. The second lesson occurred three days after the first lesson at one school and seven days at the other. Students were randomly assigned to the different experimental conditions. For this purpose, they blindly drew a colored slip of paper that assigned them to an experimental group or the control group. The groups were distributed among the classrooms of the participating classes (one condition for each room). The survey was administered by one of the authors and two trained research assistants. The instructions and the procedure of data collection were pre-formulated and pre-structured to ensure implementation objectivity. In the experimental groups, students received an introduction that they were taking part in a reading study to explore whether they would better understand a text if they looked up unknown words in a dictionary. Control group students were told that they would be tested on their reading comprehension after reading the text. Participants in each group than received the study material: The electronic *Kindle* reader was handed out to participants in the Kindle group; the Duden group received a paper-based copy of the text and the print dictionary *Duden Universalwörterbuch.* Students in the control group only received paper-based copies of the text and could not consult a dictionary. Target words were not highlighted in the text of the control group to simulate regular reading conditions in which incidental word learning usually appears. Participants in the Kindle group and the Duden group were asked to read the text and to derive the meaning of the target words by consulting the dictionary. Participants of the control group were asked to read and understand the text and were unaware of the instruction in the experimental groups.

After participants had finished reading the study text and looking up the target words, they received a paper copy of the vocabulary test. At the same time, the study text and dictionary or *Kindle* were collected, so no reference to materials could be made when working on the test. Students sitting next to each other were given different test versions to minimize the potential for cheating. Two weeks later students took the vocabulary test again to measure long-term learning effects. At this point, data collection was carried out by the teachers of the classes during regular lessons.

Knowledge of the target words prior to the experimental procedure was not assessed because a pretest would have directed the students' attention to the target words, which in addition to dictionary use while reading would have influenced learning gains by other experimentally uncontrollable factors. Therefore, an experimental study with randomized group assignment was conducted so the value of word knowledge in the groups can be supposed to be equal across groups.

3. RESULTS

Table 2 shows the mean of raw test score value that participants achieved in the reading comprehension test and the mean number of items correctly solved by participants on the first (mR0) and second (mR1) vocabulary test.

Data analysis revealed no significant differences between the groups regarding their performance in reading comprehension (F (2, 108) = 0.58, p = .562). Also, no significant differences in reading comprehension were found between boys and girls across groups (t (109) = -0.47, p = .640) and between pupils whose family language corresponded to the language of instruction (German) and where it did not (t (106) = 1.82, p = .071). However, boys performed significantly better in the vocabulary test than girls at both measurement times (mR0: t (109) = 3.10, p < .002; mR1: t (109) = 2.93, p = .004) and pupils whose family language were identical to the language of instruction outperformed pupils whose family language differed from the language of instruction at the second vocabulary test (mR1: t (106) = 2.90, p = .002).

Table 2. Mean value of first (mR0) and second (mR1) vocabulary test and reading comprehension

Variable	*N*	mR0 *M (SD)*	mR1 *M (SD)*	Raw value reading comprehension *M (SD)*
Kindle				
Girls	22	5.36 (2.26)	4.95 (2.66)	5.32 (4.61)
Boys	11	7.00 (2.10)	5.45 (2.42)	4.18 (3.43)
Total	33	5.91 (2.31)	5.12 (2.56)	4.94 (2.24)
Duden				
Girls	14	4.43 (1.79)	4.50 (2.53)	7.21 (2.91)*
Boys	22	6.36 (2.11)	6.00 (2.35)	4.68 (3.88)*
Total	36	5.61 (2.18)	5.42 (2.53)	5.67 (3.71)
Control group				
Girls	18	4.33 (2.40)	4.72 (2.54)	5.17 (4.50)
Boys	24	5.58 (2.84)	6.63 (2.57)	6.58 (4.56)
Total	42	5.05 (2.71)	5.81 (2.70)	5.98 (4.53)
Gender				
Girls	54	4.78 (2.12)	4.76 (2.50)	5.76 (4.22)
Boys	57	6.16 (2.46)	6.16 (2.52)	5.39 (4.17)
Family language (FL)				
FL = language of instr.	70	5.71 (2.48)	6.00 (2.64)	6.04 (4.28)
FL ≠ language of instr.	38	5.05 (2.34)	4.53 (2.30)	4.53 (3.83)
n/a	3	5.49 (2.43)	5.48 (2.60)	7.67 (4.73)
Total	111	5.49 (2.43)	5.48 (2.60)	5.57 (4.18)

*Note. *There is a significant difference between the girls and boys in the Duden group regarding the mean of the raw value of reading comprehension (t (34) = -2.09, p = .044). However, the boys performed significantly better in the first vocabulary test than the girls (t (34) = 2.85, p = .007).*

Table 3 shows the correlations between the key variables. It is notable that *mR0* (r =.404) and mR1 (r = .5) correlate moderately with reading comprehension.

Table 3. Zero-order correlations

Variable	1	2	3	4	5
1.Reading comprehension	–				
2.Gender	.045	–			
3.Family language	-.174	.168	–		
4.mR0	.404**	-.285**	-.130	–	
5.mR1	.500**	-.270**	-.271**	.668**	–

*Note. *The correlation is significant on the 0.05 (two-sided) level. **The correlation is significant on the 0.01 (two-sided) level. Gender: boys = 0, girls = 1; family language: German as first language = 0; German as second language = 1*

To examine our two research questions, we fitted different generalized linear mixed models (GLMM). The models included random intercepts for subjects and items to account for the variation between individuals and items. Reading comprehension, family language, school affiliation, gender and group were used as predictors to examine their influence on the dependent variable (Posttest/Follow_up). Posttest refers to the performance on each item of the vocabulary test immediately after reading the text; follow-up refers to the performance on each item of the vocabulary test taken two weeks after the Posttest. We calculated the GLMM with the R package lme4. Regarding the Posttest, the following model using reading comprehension, gender, family language and group as predictors showed best model fit (AIC = 1594.9, BIC = 1636.3): posttest ~ 1 + reading_comprehension + gender + family_language + group + (1|person)+(1|item).

The estimated variances of the random intercepts were 0.2838 (*SD* = .5327) for individuals and 0.6384 (*SD* = .7990) for items. These results show considerable variation between the individuals and the items analyzed. To quantify the explanatory power of the model, we calculated pseudo R^2 values following Nakagawa and Schielzeth (2013). The marginal R^2, representing the variance explained by fixed effects, was 0.075, while the conditional R^2, including both fixed and random effects, was 0.277. These values suggest that the fixed effects account for a relatively small portion of the variance in posttest scores, with the inclusion of random effects substantially increasing the model's explanatory power.

The fixed effects estimates for the posttest-scores are presented in Table 4. Reading comprehension had a highly significant effect on the posttest-scores. The model predicts that the better the reading comprehension, the higher the probability of solving an item of the vocabulary test correctly. Gender was also a strong predictor in the model with being female having a negative effect. In addition, group membership had a significant effect: Compared to the Kindle group, being in the Duden and control groups significantly reduces the probability of solving an item correctly. This means, that the probability to solve an item correctly is the highest when using the look-up function while reading on the *Kindle*. Looking-up unknown words in the print dictionary *Duden* results in a lower probability of correct item solution than using the kindle. A calculation of the model with different group names showed no

significant differences between the Duden group and the control group (*B* = 0.23282, *SE* = 0.19254, *z* = 1.209, *p* = .227). Thus, the research hypothesis, participants using a dictionary are expected to perform better in the vocabulary test than the control group is falsified.

Table 4. Estimated effects posttest

	B	*SE*	*z*	*p*
Intercept	- .06617	.31826	-0.208	
Reading comprehension	.11193	.02037	5.494	<. 001
Gender	-.71322	.17110	-4.168	<. 001
Family Language	-.04308	.17481	-0.246	.805
Duden	-.41967	.21154	-1.984	.047
Control group	-.65250	.20679	-3.155	.001

The best fitting model for the follow-up-test is the model using reading comprehension, gender and family language as predictors (AIC = 1531.4, BIC = 1562.4, see table 5): follow_up ~ 1 + reading_comprehension + gender + family_language+(1|person)+(1|item).

The estimated variances of the random intercepts were 0.4128 (*SD* = .6425) for individuals and 0.8478 (*SD* = .9208) for items. Calculations of pseudo R^2 values resulted in a marginal R^2 of 0.094 and a conditional R^2 of 0.345.

Unlike the posttest-model, group does not contribute to a better fit of the model. This indicates that there were no long-term learning effects of looking-up unknown words while reading.

Table 5. Estimated effects follow-up

	B	*SE*	z	p
Intercept	- .54361	.32328	-1.682	0.093
Reading comprehension	.13562	.02245	6.041	<. 001
Gender	-.62008	.18315	-3.386	<. 001
Family Language	-.34840	.19332	-1.802	.072

4. DISCUSSION

This study aimed to examine the short- and long-term effects of dictionary use while reading on the acquisition of word meaning. Another objective was to explore whether there are differences in word meaning acquisition that are related to the dictionary used (electronic/print dictionary). Results show that participants who read and looked up words on the electronic reading device *Kindle* performed significantly better in the first vocabulary test compared to both, participants who looked up words in the print dictionary and the control group who could not consult any dictionary. Thus, a direct learning effect from the use of dictionary definitions while reading on word knowledge can be stated for the Kindle group, but not for the Duden

group. The research hypothesis that using dictionary definitions generally supports the acquisition of word knowledge cannot be accepted, since no significant differences in word learning were found between the Duden group and the control group. However, empirical results support the hypothesis that students using an electronic reading device perform better in the vocabulary test than students using a print dictionary. In contrast, no significant effect was found for group for long-term learning effects: 14 days after consulting the dictionary there was no statistically significant difference whether dictionary definitions were used or not.

The results of the study are consistent with those of Reinking and Rickman (1990), in which statistically significant differences were found between the computer using group and the group looking up words in a print dictionary. In this context, it remains unclear which factors are responsible for the learning gains when using electronic reading devices. A possible explanation was mentioned above: Due to lower demands on working memory when reading and looking up words on electronic readers, students are more successful in deriving the meaning of unknown words by using a dictionary. It seems that they can focus exclusively on the derivation of word meaning instead of expending cognitive resources on mentally representing the context, which is necessitated by the competing demands of looking up in a print dictionary. An alternative possibility is that the effects in the present study arose from motivational aspects associated with using electronic reading devices at school.

It also remains unclear why the Duden group did not show greater benefit from the dictionary definitions compared to the control group. One possible cause may be that participants had less routine using a print dictionary. This may have led to lengthy lookups, possibly ending in not finding the relevant dictionary entry or giving up the dictionary consultation. A further consideration is whether the process of deriving the meaning of unknown words by using print dictionaries while reading is cognitively too demanding for seventh-grade students. The statistically non-significant differences between the Duden group and control group may also be explained by the control group being more successful with the strategy of deriving the meaning of unknown words by using the context or by incidental word learning. On the contrary, the assumption that using print dictionaries is more effective for word learning, as stated in the context of the *involvement load hypothesis*, cannot be confirmed based on the present study. According to the involvement load hypothesis, the Duden group could have been expected to perform better than the Kindle group and the control group due to the greater involvement caused by looking up words in the print dictionary.

Unsurprisingly, the study did not find any statistically significant effect of group membership in terms of long-term learning effects, since no further vocabulary exercises took place between the first and second vocabulary test. Vocabulary exercises following the derivation of word meaning while reading by using dictionaries would be necessary to establish decontextualized word knowledge.

The study is limited primarily by the fact that the students' skills in looking-up words in a print dictionary and the actual look-up behavior while reading (e.g. the

number of look-ups) were not recorded. In a replication study it would therefore make sense to collect data on the subjects' reference skills as well as on the look-up-behavior during the experiment to include them as variables in data analysis. However, as print dictionaries are likely to disappear from classrooms, future research should focus on strategies that students use when consulting electronic dictionaries or AI-tools like Chat-GPT while reading.

4.1 Implications

The results of this study indicate that seventh graders demonstrate enhanced acquisition of word meaning when they utilize the integrated lookup function in electronic reading devices. Compared to other methods, such as looking up words in print dictionaries or engaging in incidental word learning, using integrated dictionaries appears to be a more effective approach. A rejection of dictionary use while reading or the refusal of dictionary definitions as a useful learning tool for students as demonstrated for example by Beck et al. (2002), could not be supported by the empirical findings, at least for secondary school students. The recommendation to look up unknown words while reading should nevertheless be met with skepticism, since its usefulness for students' acquisition of word knowledge seems to be linked to the dictionary used (print vs. electronic) and to the level of reading comprehension. Furthermore, it seems reasonable to hypothesize that students' prior experience with dictionary definitions for word meaning acquisition may influence the outcome. It is therefore recommended that teachers foster students' strategies for utilizing dictionary definitions to optimize the benefits derived from this resource. As learning effects seem to be only short term, it is essential that educators plan subsequent exercises regarding the words looked up in the classroom to facilitate their long-term integration into students' vocabulary.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Reinold Funke for his valuable comments on the manuscript and his assistance with the statistical analysis. We also thank Peter Edelsbrunner for his support in analyzing the model fit of the vocabulary test.

REFERENCES

Alharbi, M. A. (2016). Using different types of dictionaries for improving EFL reading comprehension and vocabulary learning. *The JALT CALL Journal*, *12*(2), 123–149. https://doi.org/10.29140/jaltcall.v12n2.204

Aust, R., Kelley, M. J., & Roby, W. (1993). The use of hyper-reference and conventional dictionaries. *Educational Technology Research and Development*, *41*(4), 63–73.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life. Robust vocabulary instruction*. Guilford Press.

Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions learning the meanings of words: an instance-based learning approach. *Discourse Processes*, *45*(2), 122–159. https://doi.org/10.1080/01638530701792826

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*(4), 671–681. https://doi.org/10.1037/0022-0663.96.4.671

Carnine, D., Kameenui E. J., & Coyle, G. (1984). Utilization of contextual information in determining the meaning of unfamiliar words. *Reading Research Quarterly, 19*(2), 188–204.

Chen, Y. (2010). Dictionary use and EFL Learning. A contrastive study of pocket electronic dictionaries and paper dictionaries. *International Journal of Lexicography*, *23*(3), 275–306. https://doi.org/10.1093/ijl/ecq013

Dambeck, H. (2008, April 17). Klimawandel in den Alpen. Kunstschnee mit Kollateralschäden. [Climate change in the Alps. Artificial snow and its collateral damage]. *Spiegel Online*. http://www.spiegel.de/wissenschaft/natur/klimawandel-in-den-alpen-kunstschnee-mit-kollateralschaeden-a-547765.html

Diehr, B., Gieseler, Ch., & Kassel J. P. (2013). Lesen mit portablen elektronischen Wörterbüchern—Ausgewählte Ergebnisse der MobiDic-Studie. [Reading with portable electronic dictionaries—Selected results of the MobiDic study]. *Zeitschrift für Fremdsprachenforschung*, *24*(2), 229–261.

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation* (5th edition). Springer.

DUW[7] = *Duden, Deutsches Universalwörterbuch*. 7., überarbeitete und erweiterte Auflage, hrsg. von der Dudenredaktion. Dudenverlag, 2011.

Dziemianko, A. (2010). Paper or electronic? The role of dictionary form in language reception, production and the retention of meaning and collocations. *International Journal of Lexicography*, *3*(3), 257–273. https://doi.org/10.1093/ijl/ecp040

Dziemianko, A. (2017). Dictionary form in decoding, encoding and retention: Further insights. *ReCALL*, *29*(3), 335–356. https://doi.org/10.1017/S0958344017000131

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fischer, U. (1994). Learning words from context and dictionaries: An experimental comparison. *Applied Psycholinguistics*, *15*(4), 551–574. https://doi.org/10.1017/S0142716400006901

Flynn, M. H. (2007). Electronic dictionaries, printed dictionaries and no dictionaries: the effects on vocabulary knowledge and reading comprehension. [Master´s thesis, University of Birmingham].

Fukkink, R. G. (2005). Deriving word meaning from written context: a process analysis. *Learning and Instruction*, *15*(1), 23–43. https://doi.org/10.1016/j.learninstruc.2004.12.002

Fukkink, R. G., Blok, H., & de Glopper, K. (2001). Deriving word meaning from written context: a multicomponential skill. *Language Learning, 51*(3), 477–496. https://doi.org/10.1111/0023-8333.00162

Graves, M. F., Schneider, S., & Ringstaff, C. (2018). Empowering students with word-learning strategies: Teach a child to fish. *The Reading Teacher*, *71*(5), 533–543. https://doi.org/10.1002/trtr.1644

Herman, P. A., Anderson, R. C., Pearson, P. D., & Nagy, W. E. (1987). Incidental acquisition of word meaning from expositions with varied text features. *Reading Research Quarterly*, *22*(3), 263–284. https://doi.org/10.2307/747745

Hodson, J. (2021). Wörterbuchnutzung zur Erschließung der Bedeutung unbekannter Wörter beim Lesen: Mediale Effekte und kognitive Prozesse. SLLD-B. https://doi.org/10.46586/SLLD.198.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, *51*(3), 539–558. https://doi.org/10.1111/0023-8333.00164

Jenkins, J., Stein, M., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal, 21*(4), 767–787. https://doi:10.2307/1163000

Jenkins, J. R., Matlock, B., & Slocum, T. A. (1989). Two approaches to vocabulary instruction: the teaching of individual word meanings and practice in deriving word meanings from context. *Reading Research Quarterly*, *24*(2), 215–235.

Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal*, *78*(3), 285–299. https://doi.org/10.1111/j.1540-4781.1994.tb02043.x

Kobayashi, C. (2007). Comparing electronic and printed dictionaries: Their effects on lexical processing strategy use, word retention and reading comprehension. In K. Bradford-Watts (Ed.): *JALT 2006 Conference Proceedings* (pp. 657–671).

Koyama, T., & Takeuchi, O. (2003). Printed dictionaries vs. electronic dictionaries. A pilot-study on how Japanese EFL learners differ in using dictionaries. *Language Education & Technology*, *40*, 61–79. https://doi.org/10.24539/let.40.0_61

Koyama, T., & Takeuchi, O. (2004a). How look-up frequency affects EFL-learning? An empirical study on the use of handheld-electronic dictionaries. *Proceedings of CLaSIC 2004*, 1018–1024.

Koyama, T., & Takeuchi, O. (2004b). Comparing electronic and printed dictionaries: How the difference affected EFL learning. *JACET Bulletin*, *38*, 33–46.

Koyama, T. & Takeuchi, O., (2007). Does look-up frequency help reading comprehension of EFL learners? Two empirical studies of electronic dictionaries. *CALICO Journal*, *25*(1), 110–125.

Laufer, B. (2000). Electronic dictionaries and incidental vocabulary acquisition: Does technology make a difference? In U. Heid, S. Evert, E. Lehmann & C. Rohrer (Eds.), *EURALEX* (pp. 849–854). Universität Stuttgart.

Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Applied Linguistics*, *22*(1), 1–16. https://doi.org/10.1177/1362168811412019

Lew, R., Mickiewicz, A., & de Schryver, G.-M. (2014). Dictionary users in the digital revolution. *International Journal of Lexicography*, *27*(4), 341–359. https://doi.org/10.1093/ijl/ecu011

Liu, T-Ch., & Lin, P.-H. (2011). What comes with technological convenience? Exploring the behaviors and performances of learning with computer-mediated dictionaries. *Computers in Human Behavior*, *27*(1), 373–383. https://doi.org/10.1016/j.chb.2010.08.016

Luppescu, S., & Day, R. R. (1993). Reading, dictionaries, and vocabulary learning. *Language Learning*, *43*(2), 263–287. https://doi.org/10.1111/j.1467-1770.1992.tb00717.x

McCreary, D. R., & Dolezal, F. T. (1999). A study of dictionary use by ESL students in an American university. *International Journal of Lexicography*, *12*(2), 107–145. https://doi.org/10.1093/ijl/12.2.107

McKeown, M. G. (1993). Creating effective definitions for young word learners. *Reading Research Quarterly, 28*(1), 16–31.

McKeown, M. G. (1985). The acquisition of word meaning from context by children of high and low ability. *Reading Research Quarterly, 20*(4), 482–496.

Merisuo-Storm, T., & Soininen, M. (2010). Primary school pupils deriving word meaning from written context. *Procedia Social and Behavioral Sciences*, *2*(2), 1625–1629. https://doi.org/10.1016/j.sbspro.2010.03.248

Miller, G. A., & Gildea, P. M. (1985). How to misread a dictionary. *AILA bulletin*, 13–26.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, *20*(2), 233–253.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, *24*(2), 237–270. https://doi.org/10.2307/1162893

Nesi, H. (2000). On screen or in print? Students' use of a learner's dictionary on CD-ROM and in book form. In P. A. Howarth & R. Herington (Eds.), *EAP learning technologies* (pp. 106–114). Leeds University Press.

Nist, S. L., & Olejnik, S. (1995). The role of context and dictionary definitions on varying levels of word knowledge. *Reading Research Quarterly*, *30*(2), 172–193.

Osaki, S., Ochiai, N., Iso, T., & Aizawa, K. (2003). Electronic dictionary vs. paper dictionary: accessing the appropriate meaning, reading comprehension and retention. In M. Murata, S. Yamada & Y. Tono (Eds.): *Dictionaries and language learning: how can dictionaries help human and machine learning? Papers submitted to the third ASIALEX biennial international conference* (pp. 205–212). Asian Association for Lexicography.

Reinking, D., & Rickman, S. S. (1990). The effects of computer-mediated texts on the vocabulary learning and comprehension of intermediate-grade readers. *Journal of Reading Behavior*, *22*(4), 395–411.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). Criteria for scale selection and evaluation. In J. P. Robinson, P. R. Shaver, & L. S. Whrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes* (pp. 1–16). Academic Press.

Ronald, J., & Tajino, A. (2005). A comparison of paper and electronic monolingual dictionaries: location, comprehension and retention of secondary senses. In V. B. Y, Ooi, A. Pakir, I. Talib, L. Tan, P. K. Tan & Y. Y. Tan (Eds.): *Words in Asian Cultural Contexts. Proceedings of the 4th Asialex Conference 1–3 June, 2005, Singapore* (pp. 255–261). National University of Singapore.

Schneider, W., Schlagmüller, M., & Ennemoser, M. (2007). *LGVT 6–12. Lesegeschwindigkeits- und -verständnistest für die Klassen 6-12*. Hogrefe.

Schwanenflugel, P. J., Stahl, S. A., & McFalls, E. L. (1997). Partial word knowledge and vocabulary growth during reading comprehension. *Journal of Literacy Research*, *29*(4), 531–553. https://doi.org/10.1080/10862969709547973

Scott, J. A., & Nagy, W. E. (1997). Understanding the definitions of unfamiliar verbs. *Reading Research Quarterly, 32*(2), 184–200. https://doi.org/10.1598/RRQ.32.2.4

Shizuka, T. (2003). Efficiency of information retrieval from the electronic and the printed versions of a bilingual dictionary. *Language Education & Technology, 40*, 15–33. https://doi.org/10.24539/let.40.0_15

Stahl, S. A., & Nagy, W. E. (2006). *Teaching word meanings*. Lawrence Erlbaum Associates Publishers.

Swanborn, M. S. L., & de Glopper, K. (1999). Incidental word learning while reading: a meta-analysis. *Review of Educational Research, 69*(3), 261–285. https://doi.org/10.3102/00346543069003261

Swanborn, M. S. L., & de Glopper, K. (2002). Impact of reading purpose on incidental word learning from context. *Language Learning*, *52*(1), 95–117. https://doi.org/10.1111/1467-9922.00178

Weschler, R., & Pitts, C. (2000). An experiment using electronic dictionaries with EFL students. *The Internet TESL Journal 6*.

APPENDIX

Table 1. Correlations kindle group (N = 33)

Variable	*M*	*SD*	1	2	3	4	5
Gender			–				
Family language			.027	–			
Reading comprehension	4.94	4.24	.128	-.170	–		
mR0	5.91	2.31	-.339	-.207	.412*	–	
mR1	5.12	2.56	-.094	-.267	.511**	.647**	–

*Note. *The correlation is significant on the 0.05 (two-sided) level. **The correlation is significant on the 0.01 (two-sided) level.*

Table 2. Correlations Duden group (N = 36)

Variable	*M*	*SD*	1	2	3	4	5
Gender			–				
Family language			.040	–			
Reading comprehension	5.67	3.71	.337*	-.242	–		
mR0	5.61	2.18	-.439**	-.201	.121	–	
mR1	5.42	2.53	-.293	-.283	.243	.594**	–

*Note. *The correlation is significant on the 0.05 (two-sided) level. **The correlation is significant on the 0.01 (two-sided) level.*

Table 3. Correlations control group (N = 42)

Variable	*M*	*SD*	1	2	3	4	5
Gender			–				
Family language			.348*	–			
Reading comprehension	5.98	4.53	-.156	-.113	–		
mR0	5.05	2.71	-.231	-.064	.605**	–	
mR1	5.81	2.70	-.353*	-.251	.652**	.797**	–

*Note. *The correlation is significant on the 0.05 (two-sided) level. **The correlation is significant on the 0.01 (two-sided) level.*