Reading and Writing
https://doi.org/10.1007/s11145-024-10536-5

Continuous texts or word lists? Exploring the effects and the process of repeated reading depending on the reading material and students' reading abilities

Mareike Ehlert[1] · **Jan Beck**[1] · **Natalie Förster**[2] · **Elmar Souvignier**[1]

Accepted: 14 March 2024

Abstract
Repeated reading (RR) is often recommended for promoting reading fluency, but it is unclear whether continuous texts or word lists should be used. This study tested whether the effects of RR depend on the reading material and whether these effects interact with students' prior abilities. $N = 304$ primary school students were randomly assigned to one of three conditions: (1) a repeated reading group receiving continuous texts (RR-T), (2) a repeated reading group receiving word lists (RR-W), and (3) a control group (CG). Before and after the training ($M = 13$ sessions), students' reading fluency and reading motivation were assessed. In both RR-groups, the average level and growth in (a) words read correctly per minute and (b) affect were recorded during each reading session. Multilevel modelling revealed that growth in reading fluency and reading motivation did not differ significantly between the RR groups and the CG. Process analyses showed that the number of words students read correctly per minute increased significantly per session. Students' affect remained stable at a high level throughout the reading interventions. In the RR-T, significantly more words were read correctly than in the RR-W, and this was a mediator between group and reading fluency. Students' initial levels of reading performance and reading motivation strongly predicted both process and outcome variables, but not in interaction with the reading material. Results emphasise the importance of investigating the effects of repeated reading interventions with higher dosage and of offering differentiated *methods* (instead of reading material) to struggling readers.

Keywords Reading fluency · Reading motivation · Repeated reading · Individual differences · Differentiated reading instruction

Extended author information available on the last page of the article

Published online: 11 April 2024

Introduction

Children with reading difficulties have problems recognising words accurately and quickly (Müller et al., 2020). The most recommended method to automate word recognition is repeated reading (Hattie, 2009; National Reading Panel, 2000; Therrien, 2004). In repeated reading, students read the same reading material aloud repeatedly for a defined number of times (Meyer & Felton, 1999) or until they achieve an established improvement in reading accuracy and reading speed (Samuels, 1979). Previous research has generally supported the effectiveness of repeated reading (e.g., National Reading Panel, 2000; Therrien, 2004). As an immediate effect, repeated reading thereby leads to an increased number of words read correctly and to a positive affect in students (Chard et al., 2002; Stevens et al., 2017). These positive immediate outcomes may help explain why repeated reading improves students' reading fluency and comprehension as well as their reading motivation (National Reading Panel, 2000; Strickland et al., 2013). However, some studies reported differential effects (Chard et al., 2009; Fuchs et al., 2021), suggesting that repeated reading does not improve reading performance of all students alike. Findings are further ambivalent about whether continuous texts or word lists are the reading material that should be used for students to best profit from repeated reading (Levy et al., 1997; Therrien & Kubina, 2007). This study therefore investigates whether the method of repeated reading is more effective when students read continuous texts or when they read word lists, and whether the effects of the material depend on students' prior reading skills.

Effects and process of repeated reading

The method of repeated reading is based on the idea that readers first have to gain automaticity in word recognition to be able to release cognitive resources necessary for reading comprehension (Samuels, 1979). RR methods include a variety of peer formats, such as choral reading (Hiebert, 2006), reading to themselves (Therrien & Kubina, 2007), with a fluent adult (Therrien et al., 2012), or with peers (Simmons et al., 1995). An organisational advantage of homogenous pairs is that it guarantees a high level of active reading time for every student, as children read and make progress together. The approach of homogenous peer training was therefore chosen for this study.

Reading words aloud repeatedly allows readers to better memorise the words' orthographic structure so that words are recognised faster and more accurately (Begeny et al., 2009; Chard et al., 2002; Escarpio & Barbetta, 2016). Concomitant with such an immediate increase in the number of words read correctly are readers' positive emotional experiences during reading (Stevens et al., 2017). Though rarely studied specifically for repeated reading, it has been documented that when readers feel they make immediate progress, this elicits reading enjoyment and positive task-related affect (Tulis & Fulmer, 2013).

These mechanisms can explain increases in students' reading performance, which are frequently reported as positive outcomes of repeated reading interventions when compared with a control group (National Reading Panel, 2000; $d = 0.41$). Improved reading performance typically implies reading fluency (word- and sentence-level) as the major target of repeated reading interventions, but transfer effects have also been reported for reading comprehension (Strickland et al., 2013; Therrien, 2004). This is plausible considering that when readers automate their word recognition through repeated reading, they release cognitive resources and can turn their attention to meaning (Samuels, 1979).

Because reading motivation has only rarely been assessed as an outcome variable for repeated reading interventions, effects are less clear. However, there is evidence from Blum and Koskinen (1991) who found that repeated reading not only fostered reading fluency and comprehension but also increased student motivation. Moreover, readers who improve their reading abilities typically enjoy reading more (Schiefele et al., 2012). In this regard, it is likely that repeated reading also positively influences reading motivation.

In sum, prior research supports the general effectiveness of repeated reading interventions in improving reading fluency (National Reading Panel, 2000; Therrien, 2004) and also suggests positive effects on reading motivation. These positive outcomes can be explained by the immediate effects of repeated reading, i.e., that readers read more words correctly and that they experience positive affect during repeated reading.

Interaction between students' reading abilities and reading materials

Prior research, however, also indicates that repeated reading does not improve the reading achievement of all students alike. For instance, Fuchs et al. (2021) found that while their repeated reading intervention of 38.5 treatment hours over 22 weeks had positive effects on students' reading performance compared to the control group, the lowest performing first-grade students profited more from a phonological awareness intervention than from the repeated reading intervention with continuous texts. Low-performing students are further known to profit most from methods which help them decode smaller units of words first, such as syllable-based reading (Müller et al., 2017). Moreover, Therrien's (2004) meta-analysis including readers from ages 5 to 18 years suggested that primarily students with learning disabilities improve from repeated reading. Effects of repeated reading may therefore depend on how well the method fits to a student's reading abilities (Chard et al., 2009; Therrien, 2004).

Moreover, effects may be influenced by the type of reading material that is used for repeated reading, for instance, whether children receive continuous texts or word lists (Therrien & Kubina, 2007). On the one hand, a continuous text may primarily facilitate the reading flow of higher-performing readers by embedding words in context and thus automating word recognition. This assumption is in line with top-down processes in cognitive models of reading comprehension (e.g., Kintsch, 1988), which assume that a reader's prior knowledge and context can facilitate automatic word recognition by providing expectations about the words

that are likely to appear next. Accordingly, for a small sample size of $N = 16$ students in grades 3, 4, and 5, Therrien and Kubina (2007) found that context increased both reading speed and word recognition compared to isolated words.

On the other hand, isolated words might improve automatic word recognition by allowing readers to focus on smaller components first—thus potentially suitable for low-performing readers. This idea corresponds with bottom-up processes in cognitive reading models (Kintsch, 1988), assuming that reading comprehension depends on the speed of automatic word recognition. If a reader's automatic word recognition is still very low, context may have less effect because this information is not available in a reader's working memory. It could thus be helpful for low-performing readers to focus on isolated words where no propositions need to be formed. Correspondingly, Levy et al. (1997) reported positive effects on reading fluency when one word at a time was displayed on a computer screen, though for a small sample size of $N = 28$ students in grade 4. In sum, it remains unclear whether continuous texts or word lists should be used for repeated reading, and whether the reading material's effectiveness interacts with students' prior reading skills.

Research questions

The first goal of this study was to better understand the effects of repeated reading on reading fluency and reading motivation by addressing the following research question:

(1) Do students' (word- and sentence-level) reading fluency and reading motivation differ between the two RR interventions (i.e., with continuous texts or word list) and a control group?

H_1 Students of both repeated reading interventions will achieve higher levels of reading fluency, and reading motivation than students of the control group. Differences in the variables between the RR-T and the RR-W group are explored.

Moreover, this study aimed at better understanding the learning process through which repeated reading is effective by examining the average number and the increase in words read correctly as well as the average of students' affect and increase in affect during the training. The following research questions guided this goal:

(2) Do the two RR groups differ in (i) the average number of words read correctly during each training session and the average increase in the number of words read correctly over the course of the training, and (ii) do they differ in the average affect during each training session, and the average improvement in affect over the course of the training?

H_2 As no concrete hypotheses can be derived from previous research, this research question is exploratory.

(3) Are the effects of the reading material on students' (a) reading fluency, and (b) reading motivation mediated by (i) the average number of words they read correctly during each training session, and the average increase in the number of words they read correctly over the course of the training, and (ii) the average affect during each training session, and the average improvement in affect over the course of the training?

H_3 The more words students read and the greater their learning gains during training, the higher their reading fluency after training. Effects of the repeated reading interventions (RR-T and RR-W) on reading fluency are mediated by the average number of words read correctly per minute during training and the increase in the number of words read correctly over the course of the training. Similarly, we expect that the average affect and improvement in affect during training will predict reading motivation after training and that the effects of the RR groups on reading motivation are mediated by students' average affect and improvement in affect within individual training sessions.

(4) Are the effects of the material on reading fluency, and reading motivation moderated by students' initial reading performance?

H_4 We hypothesize that the effects of the reading material are moderated by student ability, such that students with lower reading ability benefit more from reading word lists, while students with higher reading ability benefit more from reading continuous texts.

Method

Participants

The sample consisted of $N = 304$ primary school students from 16 classes of six schools in Germany (2nd grade: six classes, 35%, 3rd grade: eight classes, 53%, 4th grade: two classes, 12%). Fifty-one percent of the participants were female. Students were on average $M = 9.10$ years old ($SD = 0.81$; 2nd grade: $M = 8.26$ ($SD = 0.42$), 3rd grade: $M = 9.34$ ($SD = 0.43$), 4th grade: $M = 10.38$ ($SD = 0.38$)). Sixty-four percent of the students only spoke German at home, 30% spoke German and another language at home, and 6% only spoke another language at home. Our sample thereby corresponds to regular primary school students in Germany (Statistisches Bundesamt, 2020). The sample included whole classes of students, recruited regardless of their reading scores.

To estimate the required sample size, the reviewed literature only provides consistent evidence for Hypothesis 1, stating that repeated reading positively influences students' reading performance compared with a control group ($d = 0.41$; National Reading Panel, 2000). Therefore, we analysed power a priori for Hypothesis 1. Before data collection, we determined the power for this effect using the *mixedpower* function of the 'mixedpower' package in the statistical software R (Kumle et al., 2021; R Core Team, 2021). Results indicated that $n = 70$ students per group were required to detect an effect of 0.41 with a power of 0.80 ($\alpha = 0.05$). The preregistration, the syntax, and the results of the power analysis are provided in the OSF (https://osf.io/aqsnc/).

For the group assignment, we first ranked children for whom a declaration of consent from their parents was available at the class level based on their reading performance in the pre-test (i.e., sentence-level fluency scores). We then randomly sampled homogeneous tandems to either the CG, RR-T, or RR-W. For instance, in class 1, the pair with the highest reading ability was assigned to the CG, the pair with the second highest reading ability was assigned to the EG-T, the pair with the third highest reading ability was assigned to the EG-W and so on. In class 2, this procedure was reversed so that the pair with the highest reading ability was assigned to the EG-T, the pair with the second highest reading ability was assigned to the EG-W, the pair with the third highest reading ability was assigned to the CG and so on. We checked via a MANOVA whether participants in the assigned groups differed by chance in the covariates (i.e., word- and sentence-level reading fluency reading motivation in the pre-test; Hsu, 1989) and in demographic characteristics (gender, age, grade, family language). This was not the case: $F(7, 296) = 0.930$, $p = 0.526$, partial $\eta^2 = 0.02$, Wilk's $\Lambda = 0.96$.

Data collection followed the ethical standards provided by the German Society for Psychology (DGPs). Prior to the study, we obtained informed written consent from students' parents. All data were collected and analysed pseudonymously.

Procedure

The study followed a three-group experimental design (Fig. 1). In March 2023, a pre-test was conducted with primary school students in all groups (RR-T, RR-W, CG) to assess reading fluency and reading motivation. In the RR-T and RR-W, a repeated reading intervention was then instructed by six student research assistants who visited the schools and trained a small group of around eight students who then worked in pairs. Student research assistants were involved in both repeated reading interventions and had been qualified for the repeated reading method in a 3-day seminar at the university. They further had already gained practical experience with the method of repeated reading during their practical semester. In the control group, regular reading lessons conducted by the teacher took place. In June 2023, a post-test was conducted in all groups (RR-T, RR-W, CG), which again assessed reading fluency and reading motivation.

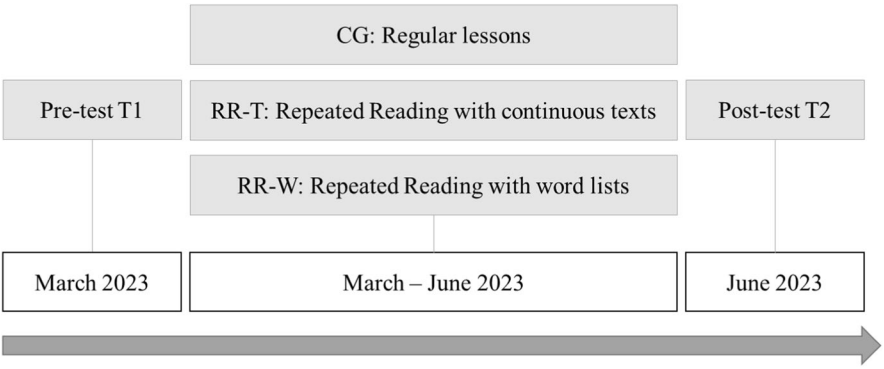

Fig. 1 Study design

Study design

While the RR-T received a repeated reading intervention with continuous texts, the RR-W received the same intervention with word lists. The continuous texts and the word lists were based on the same words so that the words read in both groups were of identical difficulty. We first randomly arranged the first half and then the second half of the text as a word list, so that the distribution of difficult words was similar. All texts and word lists ranged from 141 to 157 words and had the same difficulty level, as indicated by common readability indices (e.g., LIX: < 40; Lenhard & Lenhard, 2014–2022). In both repeated reading interventions, the procedure was the same: Two students with similar reading levels worked in pairs. Their roles were named athlete and coach. In the first reading, the athlete read a word list or text out loud correctly and as fast as possible for 1 min. Meanwhile, the coach read quietly in his or her booklet and underlined the mistakes that the athlete made. After the reading, both children counted the total number of words read correctly they had read in 1 min. The children then switched roles and the new athlete now read the same list or text aloud. This procedure was then repeated until both children had read the same material out loud twice. The research assistants supervised that the children adhered to the general principles of repeated reading (reading aloud, reading repeatedly, taking over the roles of the "coach" and the "athlete") and measured the time of each reading (1 min). They further supervised whether children entered the words read and the mistakes they made correctly. Each session (two readings) lasted approximately 15 min. On average, students in both EGs participated in $M = 12.89$ ($SD = 3.71$, $Mdn = 14$, $Min = 4$, $Max = 18$) training sessions over the course of three and a half months. The complete reading material for both EGs is provided in the OSF (https://osf.io/aqsnc/) (Fig. 2).

Fig. 2 Study material

Study material

Note. In the table, the number of words read within 1 min ("gelesene Wörter") and the reading errors ("Lese-Fehler") were recorded for each reading. These two were used to calculate the number of words read correctly per minute ("Wörter minus Lese-Fehler").

The repeated reading sessions were carried out at the back of the class. The research assistants reported what the CG was doing while the additional repeated reading took place in the intervention groups. Unfortunately, 42% of this data is missing, but the available data indicate that the CG had reading lessons (28%), maths lessons (20%), science lessons (18%), physical education lessons (12%), free-work periods in the afternoon (10%), and other subjects such as English, arts or lessons with a substitute teacher (12%).

Measures

Students' word-level reading fluency was assessed with the subtest for word-level reading of the standardised test ELFE II (Lenhard et al., 2018), in which words have to be matched with the correct image within 3 min. Again, a parallel form was used in the post-test. Parallel test reliability was $r = 0.87$ in our data. Percentages of accurate answers were 97% accuracy for T1 and 98% accuracy for T2.

Sentence-level reading fluency was assessed using the standardised test Salzburg Reading Screening 2–9 (SLS 2–9; Wimmer & Mayringer, 2014). Within 3 min, students have to decide for as many sentences as possible whether a sentence is correct or incorrect (e.g., "Trees can talk"). In the post-test, a parallel form was used, for which parallel test reliabilities of $r = 0.89$ were computed. Reading accuracy in the sentence-level reading test was high ($M_{T1} = 1.44$ wrong sentences ($SD_{T1} = 3.23$), $M_{T2} = 1.03$ wrong sentences ($SD_{T2} = 2.56$), indicating that children had no problems to comprehend the simple semantic content.

To assess students' intrinsic reading motivation, an established scale of five items was used at both measurement points (McElvany et al., 2008). Children rated their agreement on a four-point Likert scale (e.g., "I like to read."; 1 = not at all true, 2 = less true, 3 = rather true, and 4 = completely true). Cronbach's Alpha was $\alpha = 0.78$ in the pre-test and $\alpha = 0.84$ in the post-test.

Process variables were assessed in the RR-T and the RR-W immediately after students read a continuous text respectively a word list. Student research assistants instructed the students to count the number of words read and the reading mistakes of the first as well as of the second reading. Moreover, students rated their affect by answering the question "How do you now feel after this exercise?" using five emoticons, ranging from a very sad to a very happy emoticon.

Statistical analyses

All dependent variables (word-level reading fluency, sentence-level reading fluency, reading motivation) showed a high intra-class correlation ($18.8\% \leq ICC1 \leq 27.6\%$), indicating that a considerable part of the total variation is due to differences between classes. We thus decided to analyse group differences (research question 1) with multilevel models using the *nlme* package (3.1–157, Pinheiro et al., 2022) in R (4.1.1, R Core Team, 2021). For easier interpretation of group effects, we created a group variable using Helmert contrast coding to test both EGs against the CG. We fitted three multilevel models for (1) word-level reading fluency, (2) sentence-level reading fluency, and (3) reading motivation using an ANCOVA approach. This means the respective T2 score was predicted by the T1 score plus the contrast variable for the group. Students' T2 score was nested within the respective class of the students.

Before testing research questions 2, 3, and 4, we estimated students' intercepts and slopes of the words they read correctly and the affect they experienced during the reading training. To that end, we fitted two unconditional growth models using the *nlme* package in R. The first model contained the words read correctly as the dependent variable and the second model contained students' affect as the dependent variable. The session number as the 'time' variable thereby served as a predictor in the two growth models. Each measurement was nested within a student to identify each student's random effects. These random effects were used subsequently in the path models as variables of intercept and slope of words read correctly per minute and affect, respectively.

For research questions 2, 3, and 4, three path models using the *lavaan* package (Rosseel, 2012), in which (1) the T2 score of word-level reading fluency, (2) the T2 score of sentence-level reading fluency, and (3) the T2 score of reading motivation were predicted by its corresponding T1 scores. The group variable entered the model as a dummy-coded variable (RR-T vs. RR-W) and was fitted as a predictor of the respective T2 score. The previously estimated individual intercepts and slopes of words read correctly (Model 1 and 2) or of affect (Model 3) were integrated as mediators between the group and the T2 score. We additionally included the product of students' T1 scores and the group to explore whether the effects of the material on reading fluency and reading motivation were moderated by students' initial reading performance (research question 4). The total number of continuous texts respectively word lists children read during the interventions were included as predictors to account for potential confounding effects of dosage. To test hypothesis 3, we ran an additional model in which the indirect effects of group (RR-T vs. RR-W) via the intercept and slope were modelled on T2 for the respective dependent variable. Missing data for the dependent variables ranged from 0.03% to 4.61%. For the first ten sessions, 9.21% of the data were missing for the words read correctly (i.e., the average number of words in the two readings) and 14.14% of the data were missing for affect. Missing data in the path models were estimated using the full information maximum likelihood procedure (e.g., McArdle, 1994). As a robustness check, we additionally ran all path models with bootstrap standard errors by specifying "se = bootstrap" in the fitting function. Because the results were almost the same, we report the default standard errors based on the Delta method (Dorfman, 1938) and provide the results from the robustness check based on bootstrap standard errors in the supplemental material in the OSF (https://osf.io/aqsnc/). We also estimated a model that controlled for students' grade. It displayed similar results. Only two small differences were found: (1) The p-value for the relation between dosage and word-level fluency T2 was now $p = 0.162$, and (2) grade positively predicted the intercept of slope ($b = 0.482$; $p < .01$) and negatively predicted the slope of affect ($b = -0.03$; $p < .01$). Due to the small sample size for children in grade 4, we opted for the parsimonious models without grade level as a predictor. The model controlling for students' grade is provided in the OSF (https://osf.io/aqsnc/).

Results

Effects of repeated reading (research question 1)

Means and standard deviations for the dependent variables (word-level reading fluency, sentence-level reading fluency, and reading motivation) are provided for each group and measurement in Table 1.

Multilevel models revealed no significant group differences between the two EGs and the CG, neither for word-level reading fluency ($b = -0.86$, $SE = 0.84$, $p = .314$), nor for sentence-level reading fluency ($b = 0.24$, $SE = 0.68$, $p = .715$) nor for reading motivation ($b = -0.04$, $SE = 0.06$, $p = .199$).

Table 1 Means and standard deviations for all dependent variables in the pre- and post-test

	Continuous texts (RR-T, $n = 74$)		Word lists (RR-W, $n = 74$)		Control group (CG, $n = 156$)	
	T1	T2	T1	T2	T1	T2
Word-level reading fluency	43.40 (13.09)	47.06 (13.68)	43.17 (13.93)	48.19 (14.19)	43.01 (13.16)	47.93 (13.52)
T scores for word-level reading fluency	53.1 (10.2)	54.39 (10.67)	53.19 (10.2)	54.48 (10.65)	53.3 (9.9)	54.62 (10.56)
Sentence-level reading fluency	29.15 (11.91)	33.71 (11.68)	30.28 (10.81)	36.03 (10.43)	30.50 (11.01)	34.72 (11.62)
T scores for sentence-level reading fluency	44.1 (10.18)	51.62 (11.78)	44.89 (10.17)	53.46 (11.69)	44.91 (10.29)	52.77 (11.73)
Reading motivation	3.41 (0.62)	3.45 (0.62)	3.37 (0.61)	3.40 (0.66)	3.48 (0.59)	3.43 (0.67)

Standard deviations are given in parentheses. Word-level fluency test (ELFE II): Max = 75 points. Sentence-level reading fluency test (SLS): Max = 100 sentences. Reading motivation: Likert Scale 1–4

Table 2 Means and standard deviations for all process variables in the RR-T and the RR-W

	Continuous texts (RR-T, $n=74$)	Word lists (RR-W, $n=74$)
Words read correctly per minute: Intercept	90.8 (30.6)	71.6 (24.5)
Words read correctly per minute: Slope	0.91 (0.94)	0.85 (1.51)
Affect: Intercept	3.89 (0.91)	3.86 (0.95)
Affect: Slope	0.02 (0.06)	0.01 (0.06)

Standard deviations are given in parentheses. Words read correctly per minute: Min = 10; Max = 276. Affect: Likert Scale 1–5

Process of repeated reading (research questions 2–4)

Means and standard deviations for intercept and slope of words read correctly and of affect are provided in Table 2. On average, students in both EGs read $M=97.22$ ($SD=32.31$) words per minute per session. The mistakes students made during the sessions did not differ significantly between both EGs (RR-T: $M=1.88$, $SD=1.46$; RR-W: $M=2.16$, $SD=1.69$; $t(146)=-1.075$, $p=.284$). Students in both groups showed a mean affect of $M=3.94$ ($SD=0.96$) on a Likert scale of 1–5.

Results of path models for word-level reading fluency (Fig. 3), sentence-level reading fluency (Fig. 4) and reading motivation (Fig. 5) are provided in the respective figures. Estimates, standard errors, and *p*-values based on bootstrap analyses are additionally provided in the OSF (https://osf.io/aqsnc/).

Regarding research question 2, the path models showed that significantly more words were read correctly during a repeated reading intervention with continuous

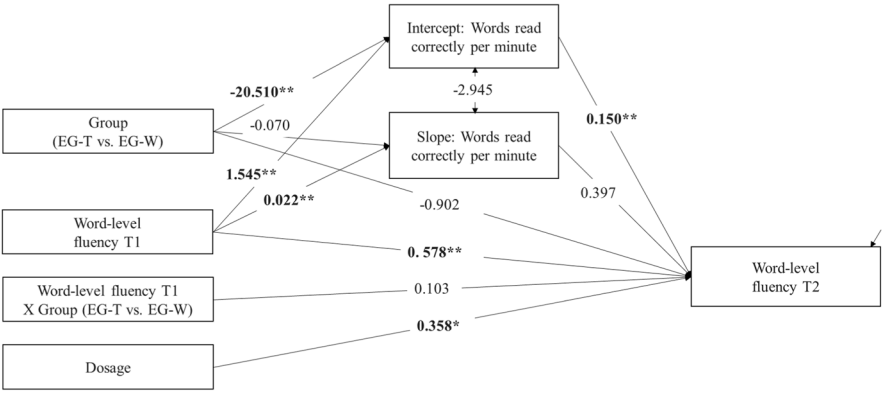

Fig. 3 Modelling the growth in word-level reading fluency. *Note*: Model fit: $\chi^2(4)=1.072$, RMSEA = 0.079, SRMR = 0.021, CFI = 0.991, TLI = 0.990. $R^2=0.814$. The coefficients displayed are unstandardised. RR-T = Repeated reading intervention with continuous texts, RR-W = Repeated reading intervention with word lists. *$p<.05$; **$p<.001$

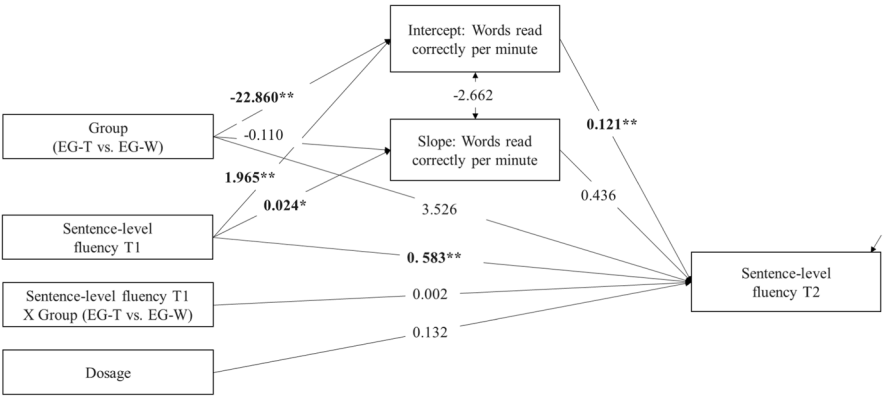

Fig. 4 Modelling the growth in sentence-level reading fluency. *Note*: Model fit: $\chi^2(4) = 1.046$, RMSEA = 0.099, SRMR = 0.024, CFI = 0.986, TLI = 0.946. $R^2 = 0.796$. The coefficients displayed are unstandardised. RR-T = Repeated reading intervention with continuous texts, RR-W = Repeated reading intervention with word lists. *$p < .05$; **$p < .001$

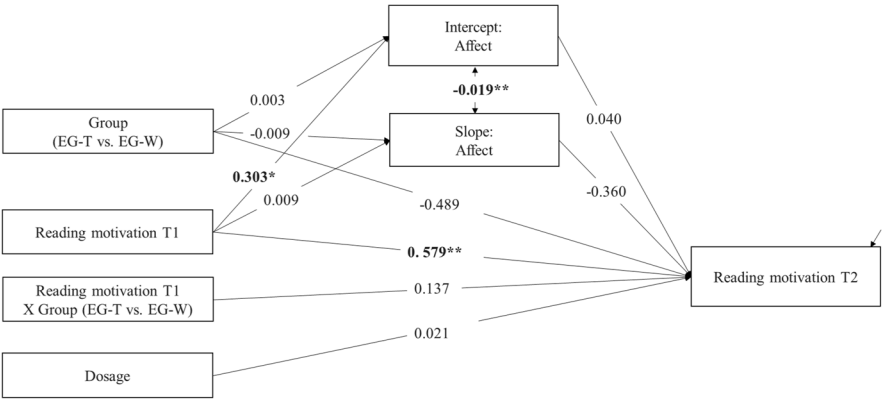

Fig. 5 Modelling the growth in reading motivation. *Note*: Model fit: $\chi^2(4) = 1.151$, RMSEA = 0.226, SRMR = 0.090, CFI = 0.707, TLI = 0.743. $R^2 = 0.454$. The coefficients displayed are unstandardised. RR-T = Repeated reading intervention with continuous texts, RR-W = Repeated reading intervention with word lists. *$p < .05$; **$p < .001$

texts than with word lists ($b = 22.860$, $SE = 2.816$, $p < .001$). However, the development from session to session (i.e., the slope) did not differ between the two conditions. We found no differences in neither the level nor the development of affect between the groups.

With regard to research question 3, we tested whether the effect of the group on word- and sentence-level reading fluency and reading motivation was mediated by the intercept and slope by modelling indirect effects. We checked the mediation despite the lack of significant effects of the group on the outcomes at T2, because a direct effect between predictor and criterion is not necessary for mediation (Rucker et al., 2011; Zhao et al., 2010). We indeed found indirect effects of group (RR-T vs.

RR-W) via the intercept of words read correctly on word-level reading fluency T2 ($b = -7.622$, $SE = 1.904$, $p < .001$, 95% CI = [−11.354, −3.890]) and on sentence-level reading fluency T2 ($b = -5.503$, $SE = 1.460$, $p < .001$, 95% CI = [−8.365, −2.641]). Students who trained with continuous texts read more words during the training than students who trained with word lists and this positively affected their word- and sentence-level reading fluency at T2. Growth (i.e., slope) in the number of words read correctly across the training was no mediator and we found no indirect effects of group on reading motivation via the intercept or slope of affect. All exact estimates, standard errors, *p*-values, and confidence intervals for indirect effects are provided in the supplemental material in the OSF (https://osf.io/aqsnc/).

Students' prior ability is a strong predictor for both process variables as well as post-test results. The effects of the material on word- and sentence-level reading fluency and reading motivation, however, were not moderated by students' initial reading performance (research question 4). Intervention dosage emerges as a significant predictor of word-level fluency T2 scores, but neither of sentence-level fluency nor of reading motivation T2 scores. Moreover, this was not the case when including grade level as a predictor.

Discussion

This study examined the effects and the process of repeated reading interventions in a regular classroom context via a three-group study design comparing regular reading instruction (CG), repeated reading with continuous texts (RR-T), and repeated reading with word lists (RR-W). In comparison with a control group, no significant effects were found for the two repeated reading interventions, neither on reading fluency nor on reading motivation. In-depth analyses comparing the two repeated reading interventions displayed that significantly more words were read correctly in the RR-T than in the RR-W. The amount of words read correctly, in turn, mediated the effect of the group on word- and sentence-level reading fluency. T1 scores of reading performance and reading motivation positively predicted both process and post-test variables, but not in interaction with the received reading material.

Regarding research question 1, the fact that we found no significant differences between the three groups was surprising, given that the additional reading practice in both interventions was provided in small groups with trained instructors and under controlled conditions. However, since the general effectiveness of repeated reading is well documented (e.g., National Reading Panel, 2000; Therrien, 2004), other plausible explanations are required. One possible reason for the lack of effects could be due to inadequate dosing of the intervention: On average, only $M = 13$ fifteen-minute sessions were conducted over 3 months. Although this is consistent with findings that reading practice should be distributed in small doses (Denton et al., 2011), a recent meta-analysis suggested that for students receiving individual reading instruction, intervention effectiveness increased after 16.8 hours (Roberts et al., 2022). For this study's interventions, this could mean that effects might become visible only if the training intensity is increased and the training is delivered over a longer period. In line with results from the path models, the significant effect

of dosage on word-level fluency scores in the post-test (Fig. 3) as well as the fact that the number of words read correctly increased significantly per session (Table 2) give more weight to this interpretation. Another explanation for missing intervention effects could be that the additional repeated reading interventions sometimes took place at perhaps inconvenient times during school (e.g., in free-work periods during the afternoon). Children might have simply been too tired at that point to fully benefit from the additional reading practice, even though this is subject to speculation.

We further examined more immediate process outcomes to gain deeper insights into the process relations of repeated reading. Regarding research question 2, descriptive statistics (Table 2) displayed that the number of words students read correctly per minute increased per session in both groups. Comparing the two groups, however, revealed that significantly more words were read correctly during the repeated reading intervention using continuous texts (RR-T) in comparison with the repeated reading intervention using word lists (RR-W). This finding underlines the idea that continuous texts facilitate a reader's reading flow so that more words are read correctly (Therrien & Kubina, 2007). Thus, one could hypothesize that, in the long run, continuous texts might be more beneficial in improving reading fluency as they lead to a higher reading dosage. In practice, however, word lists have the advantage that they can be more easily adapted to an individual student's reading level, whereas adapting texts is often more time-consuming, as they can include both too difficult and too simple words for a student. Beyond that, students' affect remained stable at a high level throughout the reading interventions and did not show significant group differences or changes over time. Based on first studies (Stevens et al., 2017), we hypothesized that repeated reading should increase positive task-related affect over time, as readers make immediate progress (Tulis & Fulmer, 2013). However, in view of the rather repetitive nature of the training procedures, the observation that students worked through the reading interventions with unchanged high affect can also be seen positively. For example, students' stable and positive affect might facilitate the interventions' implementation in regular classrooms, as teachers do not have to counteract too many adverse affectional responses (e.g., discontent, frustration). For future studies, it would be intriguing to also qualitatively assess students' affect and ask children more explicitly why they enjoyed the task.

With regard to research question 3, we found two indirect-only mediation effects, in which the number of words read correctly per minute was a mediator between group and reading fluency at T2, pointing towards a causal chain of effects (Rucker et al., 2011). In other words, students who received continuous texts were able to read more words correctly, and this, in turn, positively influenced reading fluency and comprehension at T2. The more words students read correctly, the better their reading fluency and reading comprehension developed. Because the continuous texts allowed readers to read more words correctly than words lists do, their post-test reading fluency and comprehension scores improved significantly compared with the RR-W. This finding demonstrates the mediational power of words read correctly and supports the idea that words read correctly are a relevant process mechanism in explaining increases in students' reading performance after repeated reading (Chard et al., 2002).

Considering research question 4, initial reading achievement and motivation were found to be strong predictors of reading achievement and motivation in the post-test, but not in interaction with the reading material. Here, we expected that low-performing readers would profit more from word lists, as they would not yet be able to use context information and could focus on smaller components (i.e., isolated words) first (Kintsch, 1988). Instead, readers who already performed well in the pre-tests read more words during repeated reading and also performed better in the post-tests. This phenomenon, commonly called the *Matthew Effect* (Stanovich, 1986), results in even larger differences between proficient and less proficient readers over time (Duff et al., 2015). It might therefore be more useful to provide differentiated *methods* to promote differently proficient readers instead of providing differentiated reading *materials* (Connor, 2019). For instance, there is evidence that good word recognition accuracy is a prerequisite for the development of word recognition speed (Karageorgos et al., 2020) leading to the conclusion that "attempts at promoting reading speed (…) may be a waste of time with inaccurate readers" (Juul et al., 2014; p. 1104). Thus, rather than providing repeated reading interventions with word lists for low-performing readers, they may benefit more from methods that specifically improve word recognition accuracy, such as syllable-based reading (Müller et al., 2017) or phonological awareness interventions (Fuchs et al., 2021). Practicing syllabic structures may be especially relevant for languages with orthographic consistency such as German (Ziegler & Goswami, 2005). In a larger context, the question of interaction effects between reading material and skills thus addresses the design of adaptive and differentiated reading instruction. Given our findings, it may be that differentiating between specific reading material does not make considerable differences. Thus, teachers should perhaps be more concerned with finding the appropriate *method* (e.g., syllable-based reading, repeated reading, strategy-oriented reading) for each student instead of worrying about the correct reading material to differentiate their reading instruction (Connor, 2019).

Limitations and future directions

Our study made valuable contributions to understanding the *effects* of repeated reading by including both performance and motivational variables as well as to understanding the *process* of repeated reading by recording the number of words read correctly per minute and the affect during the intervention across $M = 13$ measurement points. However, this study holds some limitations which restrict the results' generalisability to other contexts. First, there are gaps in our knowledge of the activities undertaken by the CG while the RR-groups engaged in reading practice sessions, as nearly half of this data is missing. This data also cannot not tell us about the quality or the extent of (reading) instruction that was provided in the CG, which would have been helpful to understand the standard to which the RR-groups are compared. Furthermore, even though the research assistants were tasked with externally controlling students' self-reported number of words read and their reading mistakes, they each had to supervise a group of six to eight students. This setup potentially limited their capacity to correct every

instance of miscounting or reading mistake. Moreover, the way the word lists were presented (as words separated by dots) might have hindered less proficient readers from performing well as they could already see the next word. They thus might not have been able to profit from focusing on smaller components, which may have mitigated the hypothesised interaction effects of reading proficiency with the reading material. To explore this further, future research could examine whether presenting words as separate flashcards (so that readers are not distracted by the upcoming words) facilitate reading for low-performing readers. In this regard, digital training programmes might be of particular use in presenting one word at a time and in gamifying the reading process. Finally, we suggested that the lack of effects of the two repeated reading interventions might be explained by the interventions' dosage (Miciak et al., 2018; Suggate, 2016). Investigating the long-term effects of providing continuous texts and word lists would help substantiate this interpretation and see whether continuous texts outweigh word lists over time as more words are read correctly.

To sum up, our study aimed at providing researchers and practitioners with more information on the kind of material with which repeated reading interventions should be implemented in classrooms to support children with different reading skills. We therefore compared two training conditions using either continuous texts or word lists. Our results point towards the complexity of precisely pinpointing which effects can be expected from what kind of reading material and for whom. At this point, we carefully suggest to (1) primarily focus on continuous texts when teaching repeated reading, as more words were read correctly in the RR-T than in the RR-W (Therrien & Kubina, 2007), (2) to provide less proficient readers with reading methods which specifically address lower order reading processes and word reading accuracy (e.g., syllable-based reading; Müller et al., 2017), and (3) to examine the long-term effects of more intensive repeated reading interventions (Roberts et al., 2022).

Acknowledgements We thank our student research assistants Friederike Altrogge, Charlotte Feldmann, Annika Krabbe, Sarah Maikötter, Celine Röcken and Leonie Schwinge for instructing and conducting the repeated reading interventions in the schools.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability The study's design and analysis plan were preregistered in the Open Science Framework. The preregistration, the data, reproducible codes, study material and the supplemental material are available via the following link: https://osf.io/aqsnc/.

Declarations

Conflict of interest The authors received no financial support for the research, authorship, and/or publication of this article. There are no competing interests to declare.

References

Begeny, J. C., Krouse, H. E., Ross, S. G., & Mitchell, R. C. (2009). Increasing elementary-aged students' reading fluency with small-group interventions: A comparison of repeated reading, listening passage preview, and listening only strategies. *Journal of Behavioral Education, 18*, 211–228. https://doi.org/10.1007/s10864-009-9090-9

Blum, I. H., & Koskinen, P. S. (1991). Repeated reading: A strategy for enhancing fluency and fostering expertise. *Theory into Practice, 30*(3), 195–200. https://doi.org/10.1080/00405849109543500

Chard, D. J., Ketterlin-Geller, L. R., Baker, S. K., Doabler, C., & Apichatabutra, C. (2009). Repeated reading interventions for students with learning disabilities: Status of the evidence. *Exceptional Children, 75*(3), 263–281. https://doi.org/10.1177/001440290907500301

Chard, D. J., Vaughn, S., & Tyler, B. J. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities, 35*(5), 386–406. https://doi.org/10.1177/00222194020350050101

Connor, C. M. (2019). Using technology and assessment to personalize instruction: Preventing reading problems. *Prevention Science, 20*, 89–99. https://doi.org/10.1007/s11121-017-0842-9

Denton, C. A., Barth, A. E., Fletcher, J. M., Wexler, J., Vaughn, S., Cirino, P. T., & Francis, D. J. (2011). The relations among oral and silent reading fluency and comprehension in middle school: Implications for identification and instruction of students with reading difficulties. *Scientific Studies of Reading, 15*(2), 109–135. https://doi.org/10.1080/10888431003623546

Dorfman, R. A. (1938). A note on the delta-method for finding variance formulae. *The Biometric Bulletin, 1*, 129–137.

Duff, D., Tomblin, J. B., & Catts, H. (2015). The influence of reading on vocabulary growth: A case for a Matthew effect. *Journal of Speech, Language, and Hearing Research, 58*(3), 853–864. https://doi.org/10.1044/2015_JSLHR-L-13-0310

Escarpio, R., & Barbetta, P. M. (2016). Comparison of Repeated and Non-Repeated Readings on the reading performances of students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 24*(2), 111–124. https://doi.org/10.1177/1063426615574337

Fuchs, D., Cho, E., Toste, J. R., Fuchs, L. S., Gilbert, J. K., McMaster, K. L., Svenson, E., & Thompson, A. (2021). A quasiexperimental evaluation of two versions of first-grade PALS: One with and one without repeated reading. *Exceptional Children, 87*(2), 141–162. https://doi.org/10.1177/0014402920921828

Hattie, J. A. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hiebert, E. H. (2006). Becoming fluent: What difference do texts make? In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about reading fluency* (pp. 204–226). International Reading Association.

Hsu, L. M. (1989). Random sampling, randomization, and equivalence of contrasted groups in psychotherapy outcome research. *Journal of Consulting and Clinical Psychology, 57*, 131–137. https://doi.org/10.1037/0022-006X.57.1.131

Juul, H., Poulsen, M., & Elbro, C. (2014). Separating speed from accuracy in beginning reading development. *Journal of Educational Psychology, 106*(4), 1096–1106. https://doi.org/10.1037/a0037100

Karageorgos, P., Richter, T., Haffmans, M. B., Schindler, J., & Naumann, J. (2020). The role of word-recognition accuracy in the development of word-recognition speed and reading comprehension in primary school: A longitudinal examination. *Cognitive Development*, *56,* 100949. https://doi.org/10.1016/j.cogdev.2020.100949

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163.

Kumle, L., Võ, M. L. H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods, 53*, 2528–2543. https://doi.org/10.3758/s13428-021-01546-0

Lenhard, W., & Lenhard, A. (2014–2022). *Berechnung des Lesbarkeitsindex LIX nach Björnson. [Calculation of Björnson's readability index].* http://www.psychometrica.de/lix.html. Psychometrica. https://doi.org/10.13140/RG.2.1.1512.3447

Lenhard, W., Lenhard, A., & Schneider, W. (2018). *ELFE II. Ein Leseverständnistest für Erst-bis Siebtklässler—Version II. [A reading comprehension test for 1st to 6th-graders].* Hogrefe.

Levy, B. A., Abello, B., & Lysynchuk, L. (1997). Transfer from word training to reading in context: Gains in reading fluency and comprehension. *Learning Disability Quarterly, 20*(3), 173–188. https://doi.org/10.2307/1511307

McArdle, J. J. (1994). Structural factor analysis experiments with incomplete data. *Multivariate Behavioral Research, 29*(4), 409–454. https://doi.org/10.1207/s15327906mbr2904_5

McElvany, N., Kortenbruck, M., & Becker, M. (2008). Lesekompetenz und Lesemotivation: Entwicklung und Mediation des Zusammenhangs durch Leseverhalten. *Zeitschrift Für Pädagogische Psychologie, 22*(34), 207–219. https://doi.org/10.1024/1010-0652.22.34.207

Meyer, M. S., & Felton, R. H. (1999). Repeated reading to enhance fluency: Old approaches and new directions. *Annals of Dyslexia, 49*, 283–306. https://doi.org/10.1007/s11881-999-0027-8

Miciak, J., Roberts, G., Taylor, W. P., Solis, M., Ahmed, Y., Vaughn, S., & Fletcher, J. M. (2018). The effects of one versus two years of intensive reading intervention implemented with late elementary struggling readers. *Learning Disabilities Research and Practice, 33*(1), 24–36. https://doi.org/10.1111/ldrp.12159

Müller, B., Richter, T., & Karageorgos, P. (2020). Syllable-based reading improvement: Effects on word reading and reading comprehension in grade 2. *Learning and Instruction, 66*, 101304. https://doi.org/10.1016/j.learninstruc.2020.101304

Müller, B., Richter, T., Karageorgos, P., Krawietz, S., & Ennemoser, M. (2017). Effects of a syllable-based reading intervention in poor-reading fourth graders. *Frontiers in Psychology*, *8*, 257659. https://doi.org/10.3389/fpsyg.2017.01635

National Reading Panel (U.S.) & National Institute of Child Health and Human Development (U.S.). (2000). *Report of the National Reading Panel: Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of Child Health and Human Development.

Pinheiro, J., Bates, D., & R Core Team. (2022). *Linear and Nonlinear Mixed Effects Models* (Version 3.1–157). https://cran.r-project.org/package=nlme

R Core Team. (2021). *R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing* (Version 4.1.1). https://www.R-project.org/

Roberts, G. J., Dumas, D. G., McNeish, D., & Coté, B. (2022). Understanding the dynamics of dosage response: A nonlinear meta-analysis of recent reading interventions. *Review of Educational Research, 92*(2), 209–248. https://doi.org/10.3102/00346543211051423

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: current practices and new recommendations. *Social and Personality Psychology Compass, 5*(6), 359–371. https://doi.org/10.1111/j.1751-9004.2011.00355.x

Samuels, S. J. (1979). The method of repeated readings. *The Reading Teacher, 32*, 403–408.

Schiefele, U., Schaffner, E., Möller, J., & Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behavior and competence. *Reading Research Quarterly, 47*(4), 427–463. https://doi.org/10.1002/RRQ.030

Simmons, D. C., Fuchs, L. S., Fuchs, D., Mathes, P., & Hodge, J. P. (1995). Effects of explicit teaching and peer tutoring on the reading achievement of learning-disabled and low-performing students in regular classrooms. *The Elementary School Journal, 95*(5), 387–408. https://doi.org/10.1086/461851

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360–407. https://doi.org/10.1598/RRQ.21.4.1

Statistisches Bundesamt. (2022). *Bevölkerung mit Migrationshintergrund nach Alter [Population with a migration background by age].* Retrieved March 6, 2024, https://www.bpb.de/kurz-knapp/zahlen-und-fakten/soziale-situationin-deutschland/150599/bevoelkerung-mit-migrationshintergrund-nach-alter/

Stevens, E. A., Walker, M. A., & Vaughn, S. (2017). The effects of reading fluency interventions on the reading fluency and reading comprehension performance of elementary students with learning disabilities: A synthesis of the research from 2001 to 2014. *Journal of Learning Disabilities, 50*(5), 576–590. https://doi.org/10.1177/0022219416638028

Strickland, W. D., Boon, R. T., & Spencer, V. G. (2013). The effects of Repeated Reading on the fluency and comprehension skills of elementary-age students with learning disabilities (LD), 2001–2011: A review of research and practice. *Learning Disabilities: A Contemporary Journal, 11*, 1–33.

Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities, 49*(1), 77–96. https://doi.org/10.1177/0022219414528540

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading. *Remedial and Special Education, 25*(4), 252–261. https://doi.org/10.1177/07419325040250040801

Therrien, W. J., Kirk, J. F., & Woods-Groves, S. (2012). Comparison of a reading fluency intervention with and without passage repetition on reading achievement. *Remedial and Special Education, 33*(5), 309–319. https://doi.org/10.1177/0741932511410360

Therrien, W. J., & Kubina, R. M. (2007). The importance of context in repeated reading. *Reading Improvement, 44*(4), 179.

Tulis, M., & Fulmer, S. M. (2013). Students' motivational and emotional experiences and their relationship to persistence during academic challenge in mathematics and reading. *Learning and Individual Differences, 27*, 35–46. https://doi.org/10.1016/j.lindif.2013.06.003

Wimmer, H., & Mayringer, H. (2014). *Salzburger Lese-Screening für die Schulstufen 2–9 (SLS 2–9) [Salzburg Reading Screening for grades 2–9]*. Huber.

Zhao, X., Lynch, J. G., Jr., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*(2), 197–206. https://doi.org/10.1086/651257

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3. https://doi.org/10.1037/0033-2909.131.1.3

Authors and Affiliations

Mareike Ehlert[1] · Jan Beck[1] · Natalie Förster[2] · Elmar Souvignier[1]

✉ Mareike Ehlert
mareike.ehlert@uni-muenster.de

Jan Beck
jan.beck@uni-muenster.de

Natalie Förster
natalie.foerster@uni-wuppertal.de

Elmar Souvignier
elmar.souvignier@uni-muenster.de

[1] University of Münster, Fliednerstraße 21, 48149 Münster, Germany

[2] University of Wuppertal, Wuppertal, Germany