Lexical and Sub-Lexical Reading Training on Trained and Untrained Words of L1 and L2 Students with Learning Disabilities and Behavioral Difficulties

Anne Barwasser
Kerstin Nobel
Matthias Grünke
Department of Special Education & Rehabilitation, University of Cologne, Germany

The increasing number of students who lack proficient reading skills is well known in Germany, as well as globally. Students with learning disabilities (LD), problem behavior, and German as a second language (L2) may face even greater hurdles. However, it is essential for teachers to provide adequate support to maintain equal educational opportunities. Word recognition is fundamental to reading, and thus, we designed lexical and sub-lexical unit training through a peer-tutorial repeated reading racetrack and evaluated the effects on student proficiency with trained and untrained words. A multiple-baseline design was implemented (N = 8), and the intervention took place three times a week over six weeks. The results show promising indications that the intervention is an effective tool to improve reading in a short time, yielding moderate to strong effects on trained and untrained words.

Keywords: **reading intervention, repeated reading racetracks, peer tutoring, learning and behavioral difficulties, German as a second language**

Introduction

Word Recognition Skills in the German Language

Reading is a pivotal skill for success in school across all subject areas. Upper elementary school marks a critical period for literacy development as tasks shift from learning to read to reading to learn (Reid et al., 2013). Consequently, students who continue to struggle with foundational reading skills are at risk of falling behind their peers and experiencing academic problems (Zentall, 2014), potentially limiting their future occupational choices. Research indicates that students with difficulties in word reading in the third grade are more likely to receive failing grades in school and drop out of secondary school (Brasseur-Hock et al., 2011).

Insights into Learning Disabilities is published by Learning Disabilities Worldwide (LDW). For further information about learning disabilities, LDW's many other publications and membership, please visit our website: www.ldworldwide.org.

Significant Groups with Difficulties in Reading in German Classrooms

In Germany, the literature indicates a trend toward less-proficient readers and declining reading motivation (Harju-Luukkainen et al., 2020). In fact, 25.4% of fourth-grade students do not reach the standard of reading literacy necessary for a successful transition from learning to read to learning to read (McElvany et al., 2021). Additionally, approximately one out of three students in Germany has a migration background (Federal Statistics Office, 2022), elevating the need for effective second language (L2) instruction across all classrooms. In comparison with first language students, students with an L2 background may experience difficulties in reading acquisition as well as fluency (Cop et al., 2015; Melby-Lervåg & Lervåg, 2014). Surveys indicate that although most teachers have students with L2 learning needs in their classrooms, only 30% feel they are prepared to support these students (Becker-Mrotzek et al., 2012). Statistics show that students with an L2 background in German fall behind their native-speaking peers in terms of academic success (Harju-Luukkainen et al., 2020).

Germany is striving toward a more inclusive school system. The inclusion rate rose from 19.7% in 2012 to 44.6% in 2021 in North Rhine–Westphalia (IT-NRW, 2021). Classrooms must adapt to support students with multi-lingual and German as L2 backgrounds, as well as an increasing number of students with special needs. Among students with disabilities, those with learning difficulties and behavioral problems comprise the largest groups. Research shows that students with learning disabilities (LD) and students with behavioral difficulties are especially at risk for reading difficulties (Roberts et al., 2020; Solis et al., 2012). According to the multiple-deficit or correlated liabilities model (Roberts et al., 2015; 2020), students with LD, students with an L2 background, and students with behavioral difficulties may exhibit shared risk factors, including low language proficiency, a lack of motivation, and a lack of attention. Therefore, interventions that can effectively support a variety of students with shared risk factors are becoming increasingly relevant for inclusive classrooms.

Reading Difficulties in Transparent Orthographies

The German language retains a relatively consistent letter–sound relationship, making it more transparent than, for instance, English (Gangl et al., 2018). In a relatively transparent language, poor readers often acquire high accuracy but face problems with fluency (Diamanti et al., 2018; Knoepke et al., 2014). Reading interventions in opaque languages like English often focus on accuracy. However, it is also relevant to determine which interventions are effective in improving reading speed and fluency for struggling readers of transparent languages. According to the dual route theory of reading aloud (DRT) (Coltheart et al., 1993; Coltheart, 2005), readers recognize printed words via the direct (also described as the visual or lexical) route or the indirect

(also labeled the non-lexical) route. For the indirect route, readers use a letter-to-sound approach to decode words and determine semantics. When using the direct route, readers map lexical or sub-lexical orthographic information with a representation stored in their mental lexicon to retrieve both the phonological and higher-level semantic information. When more sight words are stored in memory, greater automaticity can be achieved during the reading process, leading to gains in fluency (Ehri, 2005). Instead of focusing solely on whole-word acquisition, research also investigated the effects of supporting representations of different sub-lexical units. As Marinus et al. (2012) found, several studies that applied different approaches of letter-cluster training with onsets and rimes were unsuccessful in enhancing reading speed in struggling readers (e.g., Hintikka et al., 2008; Huemer et al., 2008; Thaler et al., 2004). Toste et al. (2019) focused on meaningful linguistic units (morphemes) and found positive effects for struggling readers in fourth and fifth grade. Mayer (2018) proposed a high-frequency training of common sub-lexical grapheme clusters for struggling readers of transparent languages to foster the storage of these patterns in the mental lexicon and facilitate the automatization of these patterns. The author focused on the most common multi-letter clusters (ind, agen, icht, etc.) across words in the German language and recommended teaching these grapheme or letter clusters in isolation first, followed by experience with numerous words containing these units. One reason behind this idea is to reduce the overall cognitive load associated with the indirect letter-to-sound approach (Mayer, 2018). In English, sub-words can be inconsistent in their mapping between spelling and phonology (e.g., _ ough in b*ough*, c*ough*, d*ough*, t*ough*, thr*ough*). However, letter clusters in German (like "ohn") will usually have identical pronunciation (S*ohn*, M*ohn*, L*ohn*) (Goswami et al., 2006). An earlier study of German letter-cluster training showed promising effects on reading fluency for struggling upper elementary students with and without disabilities (Barwasser et al., 2021a).

Methodological Framework and Study Components

Extensive research has sought to identify effective interventions to support less-proficient readers (e.g., Wanzek et al., 2010). Donegan and Wanzek (2021) recommend interventions specifically for struggling readers in upper elementary that focus on foundational skills (e.g., instruction on phonics, phonemic awareness, word reading, fluency) and comprehension (explicit vocabulary instruction, comprehension strategy instruction), as well as multi-component interventions with several focal points. Word-reading problems are the most common challenges for struggling readers in primary grades (Donegan & Wanzek, 2021; Gangl et al., 2018). Moreover, the upper elementary grades mark a transition time in which the words students are expected to read become longer and more complex (Hiebert, 2008). Harju-Luukkainen et al. (2020)

argue for using motivational components in literacy interventions, specifically in the case of struggling readers since reading can be an exhausting activity for them. Moreover, the Convention on the Rights of the Child (UNICEF, 1989) stresses the need for educating all students and developing methods to include every child. Therefore, we combined the following components (which have already been successfully evaluated) into one intervention for the present study to foster lexical and sub-lexical reading and meet the needs of the participants.

Repeated Reading through Reading Racetracks

A popular intervention to support word reading is the repeated reading procedure. During repeated reading instruction, students are asked to read a word or text more than once or until a set criterion is met (Lee & Yoon, 2015). The purpose of repeated reading is to achieve automatization, which supports support fluency and allows students to focus on higher-order processes in reading, such as comprehension (Ardoin et al., 2013). Repeated reading has been found effective in both the first language and L2 for students with and without disabilities (Gorsuch & Taguchi, 2008; Lee & Yoon, 2017; Musti-Rao et al., 2009; Stevens et al., 2017; Wu et al., 2020).

One common, playful way to incorporate repeated reading is reading racetracks. Reading racetrack interventions employ a board game that consists of empty squares containing cards with certain content (i.e., words). Several studies have demonstrated the effectiveness of these interventions to support reading in students with and without disabilities, as well as in an inclusive whole classroom setting (Barwasser et al., 2022; Barwasser et al., 2021a; Barwasser et al., 2021b; Barwasser et al., 2021c; Barwasser et al., 2021d; Sperling et al., 2019). Further, Lämsä et al. (2018) report that games have the potential to engage learners in literacy, improve basic reading skills, and increase the time students are willing to invest in training. However, according to Pikulski and Chard (2005), it is also necessary to go beyond whole-word training and attempt to achieve transfer effects on unknown words by using sub-lexical patterns (see also Huemer et al., 2008; Mayer, 2018).

Motivational Components: Peer Tutoring and Self-Graphing

It is necessary to ensure an intervention is usable by a wide range of students. Peer-tutoring interventions have proven particularly effective for reading as well as behavior (e.g., Bowman-Perrott et al., 2013; Moeyaert et al., 2021). In the peer-tutoring method, peers work together on a specific task (Dufrene et al., 2010). Lee and Yoon (2015) established that repeated reading is more effective when used in combination with peer-mediated reading. Furthermore, students with learning and behavioral difficulties can profit from peer interventions in their first language and L2 (Bowman-Perrott et al., 2013; Cole, 2014; Dunn et al., 2017). Additionally, several experiments suggest that self-graphing (tracking one's own learning progress) is a promising method to

document and visualize fluency development in reading and promote on-task behavior—thus leading to an improved learning outcome in students' first language and their L2 (Albers & Hoffmann, 2012; McKenna & Bettini, 2018). Since students with learning challenges are at risk of developing a motivational loss to learn and motivation can influence an intervention's success (Nelson & Harwood, 2010; Scanlon et al., 2017), it is important to implement additional motivational boosters. Guzman et al.'s (2018) meta-analysis of self-monitoring on K–12 students' reading demonstrated a substantial effect on reading (Tau-U = 0.79, $p < .001$). To sum up, studies have demonstrated that multi-component reading interventions are effective for all students, including those with disabilities (Afacan et al., 2017).

The Present Study and Research Aim

We conducted a thorough survey of the topic and designed an intervention that best fits students with reading difficulties and shared risk factors (LD, German L2, and behavioral problems) according to the correlated liability model (Roberts et al., 2020). In view of inclusion and the right to education for all (UNICEF, 1989), it is urgent to intensively evaluate this type of intervention and bring it to schools. So far, this multi-component intervention has been applied in only one study, with good results (Barwasser et al., 2021a). To further intensify and investigate this type of intervention, this study examines the effects of the intervention on a greater number of students, this time also including L2 students. This experiment applied a lexical and sub-lexical reading intervention using the most common German letter clusters. We evaluated the effects on trained and transfer words in low-literacy primary-school students facing additional hurdles in general learning, behavior, and L2 acquisition. Our goal was to improve the reading of trained and untrained words. Our research questions were as follows:

1) Does the use of sub-lexical units training in a multi-component intervention have positive effects on trained words in students with reading difficulties with and without LD, behavior problems, and German L2?
2) Does the use of a sub-lexical units training in a multi-component intervention have positive effects on untrained words in students with reading difficulties with and without LD, behavior problems, and German L2, and thus indicate transfer effects?
3) Do participants regard the intervention as useful, acceptable, and socially valid?

Materials and Methods

Setting and Participants

A total of 12 students from an inclusive (students with and without special needs) elementary school in an urban part of North Rhine–Westphalia, Germany, participated in the study. At the time of data collection, the ages of the six boys and six girls ranged from second to fourth grade. Seven master's students of special education conducted the data collection and intervention. Informed consent forms were sent to the legal guardian in advance.

Screenings

We conducted several tests to identify the final participants and obtain additional information about them. The purpose of these tests was to better understand the intervention's effectiveness and whether it was usable for a variety of students.

The participants were selected based on a German reading screening (*Salzburg Reading and Spelling Test*; SLRT II, Moll & Landerl, 2010). The screening identifies possible word-reading problems while students read words in one minute. The parallel test reliability coefficient is between .90 and .98, and correlations with the Salzburg Reading Screening 1–4 (Mayringer & Wimmer, 2003) range from .69 to .92. Children with a percentile rank <15 in word reading (i.e., a child with a percentile rank of 15 performs worse than 85% of the children in the comparison group) were chosen for the study.

The final sample included students with German as a first language (L1) and German as a second language (L2). Because vocabulary knowledge can influence reading, a standardized German vocabulary test (CFT; Weiß, 2006, cutoff percentile rank <15) was used to measure German vocabulary proficiency and identify it as a possible influencing variable. In the vocabulary test, which reached a good reliability of r = .87, children were shown 30 main words accompanied by a choice of five other words. The children marked which word of the five choices came closest to the main word (i.e., synonyms).

The Integrated Teacher Report Form (ITRF; Volpe et al., 2018) was used to identify disruptive behavior (e.g., "disturbs others," "has conflicts with classmates") and academically disengaged behavior (e.g., "comes to class unprepared," "does not participate in class," and "does not start working on tasks independently"). In the ITRF, teachers rate their students on eight items for each category, on a four-point scale from 0 (behavior is not problematic) to 3 (behavior is strongly problematic). In this study, we used the German short version with 16 items, due to time constraints. The cutoff for the total problem score was 13, for academically engaged behavior the cutoff was 10, and for disruptive behavior it was 5. Consistency analysis of the ITRF yielded a high value of $\alpha = .87$ to $\alpha = .91$ (Volpe et al., 2018). Since none of the subjects showed

disruptive behavior but the majority showed difficulties in academically engaged behavior, only the last appears with percentile ranks for the children in Table 1.

Final Participants

All students but Cem and Robert scored above the cutoff for academically disengaged behavior. We collected socio-demographic data on the students through a teacher questionnaire. Four students dropped out of data due to the COVID-19 quarantine. Thus, there were eight final participants. Four students had an official German LD diagnosis, meaning that these students cannot follow lessons without additional help; this diagnosis is often accompanied by a reduced IQ (70–80) (Grünke & Cavendish, 2016). In Germany, an LD diagnosis is a pedagogical (educational) diagnosis and not a medical diagnosis. The process is usually initiated when teachers notice that students are significantly behind in several subjects, and the students then undergo ability tests and an intelligence test.

Table 1. ***Describing Details of Participants***

Participants	Gender	Age	Special Needs	AEB	Reading PW (PR)	Reading W (PR)	German Vocab Proficiency (PR)	L1
Lauren	female	10	LD	10	<1	<1	12	English
Lotta	female	10	/	19	1-2	<1	72	German
Serge	male	9	/	17	<1	<1	10	Russian
Cem	male	9	LD	4	29-34	<1	2	Tunisien
Ameh	male	9	/	19	1-2	3-6	4	Marrocain
Robert	male	10	LD	9	4-5	<1	9	English
Vladi	male	8	/	13	10-13	1-2	10	Russian
Tim	male	9	LD	11.5	<1	<1	84	German

Note. Learning disabilities (LD); First Language (L1); Pseudo words (PW); Words (W); Percentile (PR); Academically Engaged Behavior (AEB); Vocabulary (Vocab)

We divided the initial 12 students into three small intervention groups. In each group, we appointed two children as pairs for peer tutoring (pairs of the same reading level). The pairs were established based on the teachers' recommendations. The pairs were equal readers, and all German L2s learned German upon entering kindergarten at age three.

Research Design

A multiple-baseline design across participants was chosen to limit possible influences from confounding variables such as history (interim events) and maturation, thus increasing the internal validity (Kazdin, 2010).

Accordingly, we used an AB (A = baseline; B = intervention) design in which interventions for students began on different days. The intervention sessions began immediately after the baseline phase, which can be seen as the "control group." The three groups started at the same time with the baseline. The intervention sessions lasted for five measurement time points for the first group, four for the second group, and six for the third group. Fourteen intervention sessions were scheduled for group 1, 15 for group 2, and 13 for group 3. Due to the COVID-19 pandemic, groups 2 and 3 only participated in 12 intervention sessions at the end. The different baseline and intervention lengths are due to experimental control. The intent is to decrease the probability of alternative explanations (maturation and history) for the intervention's effectiveness when starting on different days (see Byiers et al., 2012). We randomly assigned the groups to one of the three length conditions. Intervention sessions took place three times a week for 30 minutes over six weeks. The three intervention groups were each supervised by two master's students of special needs education. The seventh master's student performed external treatment fidelity tasks.

Procedures and Materials

Baseline

We used the baseline to assess the current state of the dependent variables, meaning that no reading intervention occurred during this time. Instead, students were assigned other tasks that had no effect on their reading performance to control the Hawthorne Effect (improving simply because they receive specific attention). In this case, students were asked to complete cognitive tasks (i.e., Which picture does not match the others? How would you logically continue the series of symbols?) in the selected pairs for the same duration as the actual intervention. After each baseline session, data was collected from each participant. The conditions were identical, despite the content, to better estimate the intervention's effectiveness.

Intervention

The intervention and each intervention session consisted of two phases, which the interventionists also supervised. The sub-lexical units used in the intervention were taken from Mayer (2018), with sub-lexical units such as "and," "ohn," and "und." In total, there were 10 sub-lexical units with associated words. These sub-lexical units are simply grapheme sequences that occur frequently in German words.

The first phase served to introduce the sub-lexical units and training words that entail these clusters. Phase (1) lasted 15 minutes. Interventionists showed the training words with the colored sub-lexical units on 8.3 × 11.7-in. cards to all students. One of the children read the respective word aloud, and another child named the corresponding sub-lexical unit. Next, all the children read the word aloud again in unison. During the first phase, students received

intensive instruction in the sub-lexical units before they would practiced alone during the second phase in their second teams.

Once the students received direct instruction in the words of the intervention session, they independently played the reading racetrack game with their peers as phase (2), which lasted 10 minutes. To reduce pressure on the students, the intervention was not timed per word on the racetrack. The racetrack playing field consisted of an 11.7 × 16.5-in. board. The resulting board game featured a racetrack with a start and finish line and small squares. The small flashcards containing the training words were distributed in these fields. Two of the 10 sub-lexical units were always included as part of a session. After the cards were distributed, the racetrack game (Figure 1) began.

After the first child rolled the dice, they moved the figure forward according to the number shown on the dice. The second child then turned the flashcard over, read the word, and named the corresponding sub-lexical unit. The first child performed a helper function. If the word was read correctly and the correct sub-lexical unit was named, the card was collected in a pile from the students who read it first. Otherwise, it was put aside. Then, the second child rolled the dice, and the game continued. The interventionists were instructed in the importance of observing the students and intervening if words were taught incorrectly. Consequently, there was targeted student monitoring. In addition, there were "help cards" for the students, which they could place on the table if they had difficulties with a word. The students were instructed to put the card on the table so an interventionist could assist them without disturbing other students.

The racetrack game was finished after 15 minutes. Each child implemented the reward system by using one self-graphing sheet for training words (Figure 2) and one for transfer words. The participants drew as many squares in a row on these sheets as the number of words they read correctly during measurements. The rows corresponded to the number of intervention sessions, and the boxes corresponded to the maximum possible number of words read correctly. A comparison with the top row from the previous intervention allowed us to observe what changes had occurred.

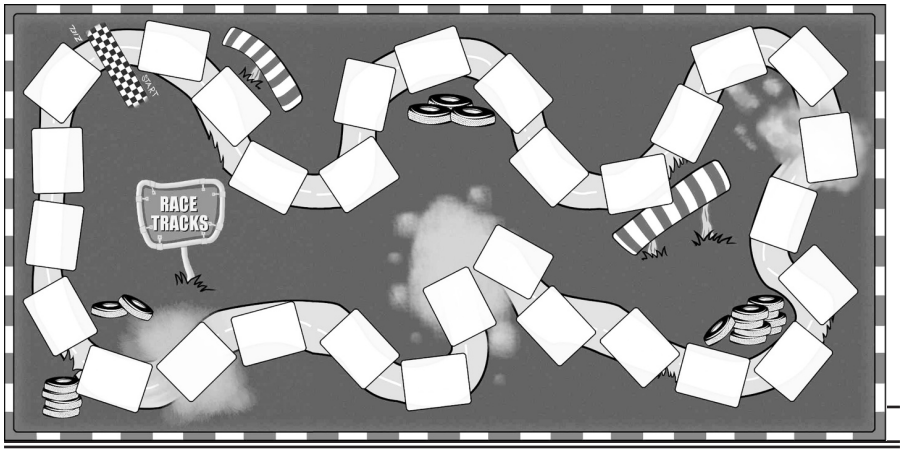

Figure 1. ***Racetrack Board***

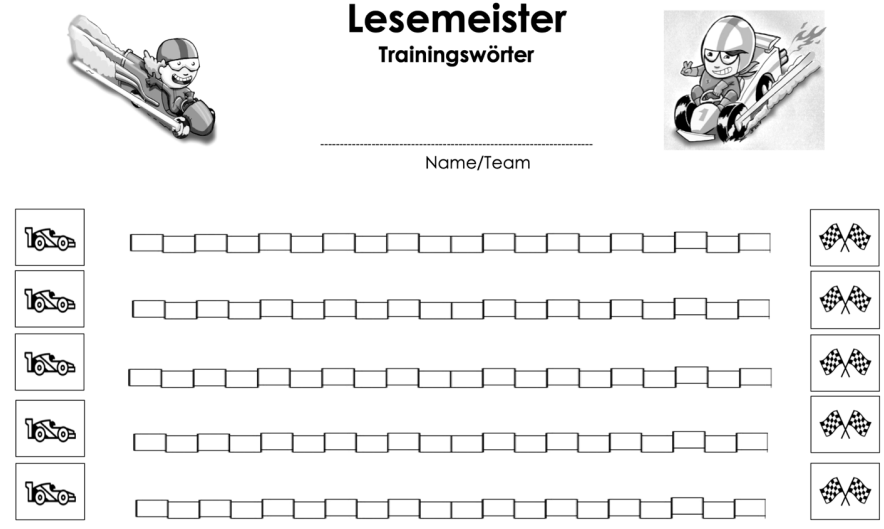

Figure 2. *Self-Graphing Sheet*

Dependent Variables and Measurements

At the end of each baseline and intervention session, possible changes in the dependent variables—(1) the *number of correctly read training words* and (2) *transfer words*—were collected, and the children were thus assessed individually. A measurement contained a total of 40 words, half of which were training words and half of which were transfer words. The sub-lexical units were equally

distributed across the words, and the two-word lists contain a similar frequency of sub-lexical units (a total of 10 sub-lexical units). The words all came from a respective pool of 70 transfer words and 70 training words (Table 2), derived from a German "Rapid Word Recognition" training by Mayer (2018) (e.g., "*ecke*"—training word: "*Hecke*," transfer word: "*Stecker*") from which 20 were always randomly included in the measurement. The overall word pool included the 10 most common grapheme clusters in the German language, and the words were based on the general German child's vocabulary. Using SPSS IBM, we calculated that the words in the assessments after each session did not differ significantly, indicating that the assessments were of equal difficulty.

Table 2. ***Training Words and the 8 Letter Clusters -Examples***

H**and**	R**echt**	L**icht**	D**ecke**
(hand)	(law)	(light)	(cover)
S**ohn**	G**ast**	**Hund**	T**orte**
(son)	(guest)	(dog)	(cake)

Note. Material taken from Mayer (2018)

The measurements were taken using a Power Point presentation (PPT); a word to be read was displayed individually on a slide for one second to check whether it was included in the child's mental lexicon (Ehri, 2005). The one-second rhythm, which is automatically set by the PowerPoint presentation and therefore does not need to be controlled externally, allowed us to assess the reading fluency of words to be captured, according to Ehri. Students were asked to correctly read the words aloud within one second of the word appearing, before the next word appeared. Then the number of correctly read training words and correctly read transfer words were documented.

Treatment Fidelity

"Treatment fidelity is commonly defined as the degree to which an intervention or program is delivered as intended" (Prowse & Nagel, 2015, p. 1). Prior to the start of the study, the master's students were intensively trained in how to conduct it. Our self-designed treatment fidelity checklist included an attendance list of the subjects and questions in six major areas: 1) external circumstances, 2) planning, 3) materials, 4) process of the intervention, 5) diagnostics and feedback, and 6) addressing student behavior. The first five areas could be answered yes/no, and the sixth area featured a value range from 0 = do not apply at all to 4 = apply completely. The interventionists always completed the sheet after each session, and for 1/3 of the intervention time, an external person was in the school to fill out the sheet for all three groups. The internal agreement between the interventionists, as well as between the interventionist and the external person, was 100%.

Social Validity

The goal of intervention is to achieve positive effects and determine how the intervention is received by those affected. Social validity is necessary to determine the usefulness and acceptance of interventions. For this purpose, we created a social validity questionnaire for students, consisting of a total of 12 items for participants to rate on a 5-point scale from 0 (do not agree at all) to 4 (completely agree). It was based on the Usage Rating Profile-Intervention by Briesch et al. (2013). The following items were rated: 1) The intervention helped me to read words correctly; 2) I think the intervention also helps other students with difficulties in reading; 3) I understood well the intention of the intervention; 4) I learned a lot during the intervention; 5) I gladly came to the intervention; 6) The intervention was fun; 7) I would like to participate again; 8) The consonant cluster helped to be a better reader; 9) Self-graphing was fun; 10) I would like to do such an intervention more often; 11) The words were difficult; and 12) I liked playing in pairs. The participants filled out the social validity questionnaire with the class teachers in the absence of the interventionists to avoid bias.

Results

Data Analysis

Data obtained in this research project will be summarized using visual analysis, descriptive statistics (Figures 3, Tables 1 and 3), and overlap indices. Furthermore, the widely used mean baseline difference (MBD) was applied (Manolov & Moeyaert, 2017). The MBD is calculated by subtracting the mean of the baseline data from the mean of the intervention data and then dividing it by the mean of the baseline data, which is then multiplied by 100. In addition, overlap measures were implemented for further analysis. The percentage of non-overlap of all pairs (NAP) indicates the percentage of data points in the intervention that are above those in the baseline. The extent to which the data do not overlap is thus determined (Parker et al., 2011). The PEM focuses on the data points that exceed the median of the baseline phase (Ma, 2006). For NAP, we applied the following cutoffs: up to .65 = weak improvement, .66 to .92 = medium improvement, .93 to 1.0 = large improvement. For PEM, we applied the following cutoffs: < .70 = non-effective treatment, .70 to .90 = moderate effects, and .90 to 1.0 = strong effects. For visual inspection as well as for data analysis, we used the statistics program R, as well as the SCAN Package for single-case data analysis, which can also test for significant results regarding the NAP.

Training Words

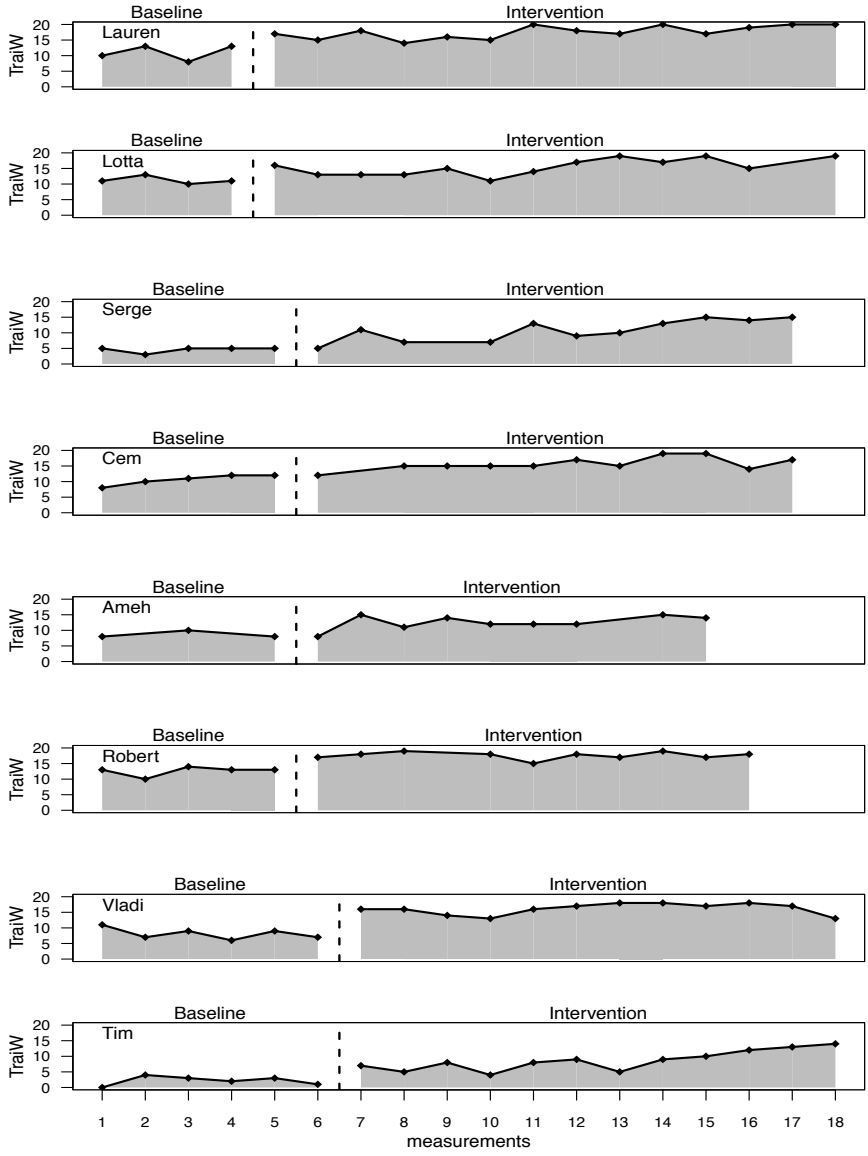

Figure 3. **Trained Words Read Correctly**

In the visual inspection, high baseline values were noted in all participants except Tim. There was an improvement from the A (baseline) to the B (intervention) phase in all children. The MBD showed the largest percentage increase for Tim (299.54%) and Serge (135.22%). Lotta (37.42%) and Robert (39.68%) improved the least. Lauren reached the maximum possible score in the B phase (20.00). Lotta, Cem, and Robert reached 19.00. The PEM results showed moderate effects for Ameh (88.89) and strong effects for Serge (90.91) and Lotta (92.31). Strong effects were also observed for Lauren, Cem, Robert, Vladi, and Tim, who attained a maximum value of 100.00. With respect to the NAP, all participants displayed a strong effect, up to an effect of 100.00. All results were considered statistically significant (<.05 to <.001).

Table 3 ***Descriptive data for training words***

Participants	N (A)	N (B)	M (A) (*SD*)	M (B) (*SD*)	Max (A)	Max (B)	MBD
Lauren	4	14	11.00(2.45)	17.57(2.06)	13.00	20.00	59.73%
Lotta	4	14	11.25(1.26)	15.46(2.63)	13.00	19.00	37.42%
Serge	5	12	4.60(0.89)	10.82(3.49)	5.00	15.00	135.22%
Cem	5	12	10.60(1.67)	15.73(2.10)	12.00	19.00	48.40%
Ameh	5	12	8.67(1.15)	12.56(2.24)	10.00	15.00	44.87%
Robert	5	12	12.60(1.52)	17.60(1.17)	14.00	19.00	39.68%
Vladi	6	12	8.17(1.83)	16.08(1.83)	11.00	18.00	96.35
Tim	6	12	2.17(1.47)	8.67(3.20)	4.00	14.00	299.54%

Note. Measurements (N); A Phase (A); B Phase (B); Mean(M); Standard Deviation (SD), Maximum Value (Max); Mean Baseline Difference (MBD)

Transfer Words

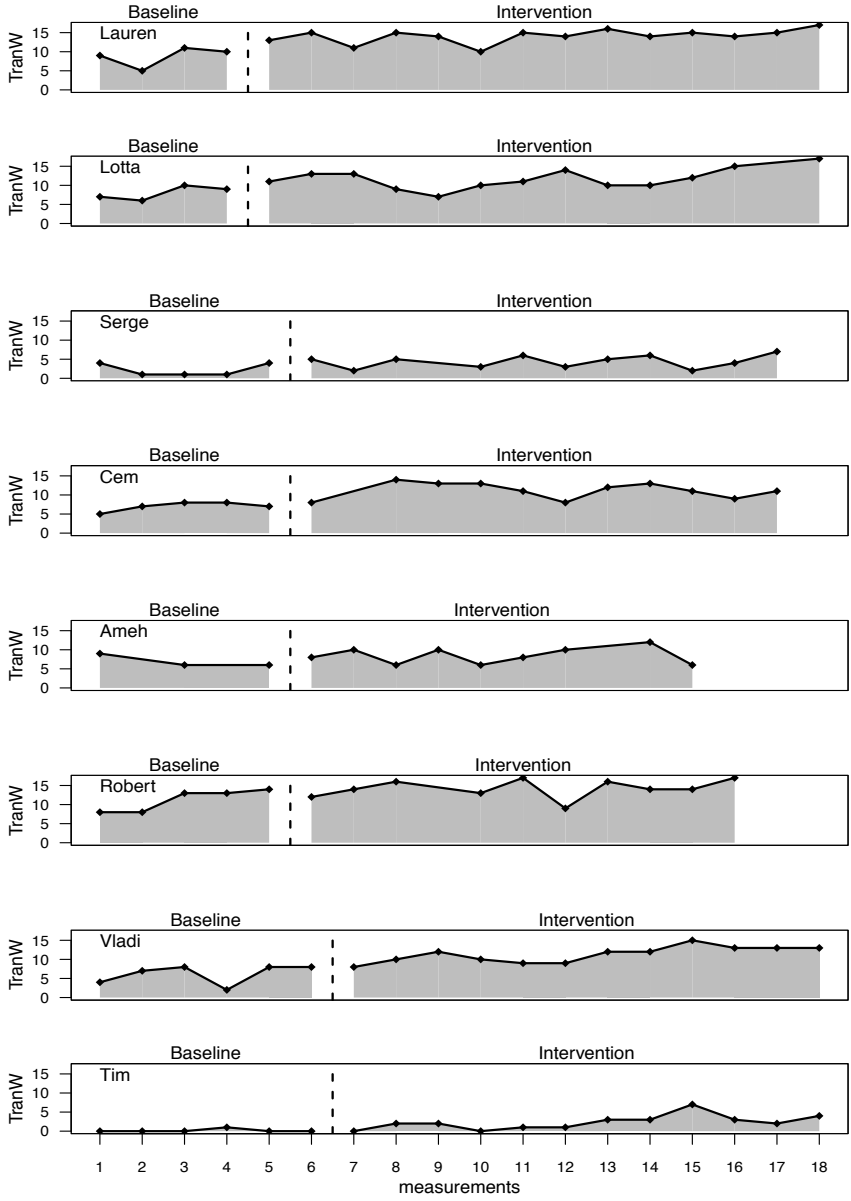

Figure 4. **Transfer Words Read Correctly**

The participants also entered the baseline phase with higher values. Serge and Tim scored the lowest values. Overall, there was an increase during the B phase for all subjects. Only Tim seemed to have difficulties with the transfer words. The MBD calculation showed the highest increase in the number of correctly read transfer words for Tim (1270.59%) and Serge (98.18%). The lowest increases were for Ameh (20.57%) and Robert (26.79%). Lauren, Lotta, and Robert reached a value of 17.00 in the B phase. The lowest values were shown here for Serge and Tim, with a value of 7. Moderate effects were found within the PEM for Robert (70.00) and Tim (83.33), while Lotta (92.31), Lauren, Serge, Cem, and Vladi (all 100.00) showed strong effects. For Ameh (66.67), the treatment did not seem to have been effective. The calculation of the NAP showed moderate effects for Ameh (70.00, $p = 1.67$), Robert (81.00, $p < .05$), Serge (84.00, $p < .05$), and Lotta and Tim (both 89.00, $p < .01$). Strong effects could be seen in Cem and Lauren (both 96.00, $p < .01$) and Vladi (98.00, $p < .001$). All results except for Ameh were statistically significant.

Table 4. ***Descriptive Data for transfer words***

	N (A)	N (B)	M (A) (*SD*)	M (B) (*SD*)	Max (A)	Max (B)	MBD
Lauren	4	14	8.75(2.63)	14.14(1.83)	11.00	17.00	61.60%
Lotta	4	14	8.00(1.83)	11.69(2.69)	10.00	17.00	46.13%
Serge	5	12	2.20(1.64)	4.36(1.69)	4.00	7.00	98.18%
Cem	5	12	7.00(1.22)	11.18(2.09)	8.00	14.00	59.71%
Ameh	5	12	7.00(1.73)	8.44(2.19)	9.00	12.00	20.57%
Robert	5	12	11.20(2.95)	14.20(2.49)	14.00	17.00	26.79%
Vladi	6	12	6.17(2.56)	11.33(2.10)	8.00	15.00	83.63%
Tim	6	12	0.17(0.41)	2.33(1.92)	1.00	7.00	1270.59%

Note. Measurements (N); A Phase (A); B Phase (B); Mean(M); Standard Deviation (SD), Maximum Value (Max); Mean Baseline Difference (MBD)

Social Validity

Overall, the children rated the intervention as highly positive, indicating that it was helpful and fun. Only Vladi rated almost all items as "partly apply." However, not all the children liked working with their peers. Serge, Cem, and Vladi rated the partner work less well. For Tim and Serge, the words were too difficult.

Table 5. ***Results of social validity questionnaire***

Items	Lauren	Lotta	Serge	Cem	Ameh	Robert	Vladi	Tim	Mean
The intervention helped me to read words correctly.	4	-	3	4	3	4	2	4	3,4
I think the intervention also helps other students with difficulties in reading.	4	-	4	4	4	4	4	4	4
I understood well the intention of the intervention.	4	-	4	4	3	4	4	4	3,9
I learned a lot during the intervention	4	-	3	4	4	4	2	4	3,6
I gladly came to the intervention.	4	-	4	3	3	4	2	4	3,4
The intervention was fun.	4	-	4	2	4	4	2	4	3,4
I would like to participate again.	4	-	4	4	3	4	2	4	3,6
The consonant cluster helped to be a better reader.	4	-	4	4	4	4	2	4	3,7
Self-graphing was fun.	4	-	4	4	4	4	3	4	3,9
I would like to do such an intervention more often.	4	-	2	4	3	4	1	4	3,1
The words were difficult.	0	-	4	2	0	0	0	0	0,9
I liked playing in pairs.	4	-	1	1	4	4	2	4	2,9

Note. Do not agree at all (0) to completely agree (4)

Discussion

Main findings

The increasing number of primary school students with reading difficulties is alarming not only in Germany but worldwide (Brasseur-Hock et al., 2011; Harju-Luukkainen et al., 2020). This challenge is compounded by motivational problems meaning that affected students usually find themselves in a vicious circle. Students who are learning an L2, have an LD, and/or have behavioral problems are among the most affected subgroups, and the gap between them and their peers is widening (Cop et al., 2015; Melby-Lervåg & Lervåg, 2014; Peterson et al., 2017). Combinations of these characteristics are highly complex and pose enormous hurdles for students and teachers. Thus, the aim of the present study was to evaluate a combined reading intervention based on theory and empirical knowledge in this field to best fit the focused student group. Thus, these findings contribute to the literature by demonstrating that this multi-component intervention can be implemented for different kinds of students, supporting the Convention on the Rights of the Child (UNICEF, 1989). Furthermore, as only one previous study had focused on this multi-component intervention for different kinds of students, the present study attempted to shrink the research gap in this area.

Overall, students achieved higher values during the baseline for both dependent variables. Nevertheless, a clear increase was observed from the A to the B phase, which was also apparent in the analysis of the data. Starting with the first dependent variable, the training words showed the smallest effect sizes for Lotta, Serge, and Ameh, who all exhibited a lack of academically engaged behavior. The problem behavior might play a role in this finding, as might Ameh's statement on the social validity questionnaire that the words were too difficult for him. Ameh, the student with the weakest increase, also performed less proficiently on the German vocabulary test (percentile rank 4) compared to Lotta and Serge. Vocabulary, which is related to reading (Lervåg & Aukrust, 2010), could therefore have played a role in this finding. Serge did not seem to enjoy partner work; he could also have been a moderator.

Regardless of personal characteristics, all subjects benefited. None of the characteristics could be conclusively identified as responsible for higher or lower performance. Instead, the findings gave slight indications. Regarding the transfer words, all subjects had great difficulties, but this was to be expected. Applying knowledge to a foreign context is more complex than memorizing sight words that have been shown to the student repeatedly (Rashotte & Torgesen, 1985). Tim, Ahme, and Robert were among the students with the lowest scores, indicating zero to moderate effects, while the others scored quite high. Serge also seemed to face difficulties; his rating indicating problems with

academically engaged behavior was one of the highest. Both Serge and Tim rated the words as "too difficult" in the intervention. Ameh, who showed moderate to zero effects, struggled with academically engaged behavior, especially regarding the application of knowledge, which requires higher concentration. Thus, a lack of academically engaged behavior could have been a factor.

According to the graph, Ameh was better at memorizing sight words; a negative trend is visible from A to B. Perhaps frustration and motivation also influenced the intervention's effectiveness, as these factors are often linked to problem behavior (Peterson et al., 2017; Roberts et al., 2015; Trzesniewski et al., 2006). Robert was diagnosed with LD and spoke English as his first language. It would be interesting to see how language background influences the learning of an L2, especially since English is a non-transparent language, while German is a transparent language (Gangl et al., 2018). However, the students' knowledge of the first language was not recorded. Lauren, who had LD and English as her first language, seemed to have fewer problems with transfer-word reading than her peers. It is possible that LD (which was present in Tim and Robert) is a significant factor. It may be challenging for these students to use acquired knowledge in an unknown context, and they may also face problems in phonological processing. Nevertheless, for Lauren, LD was not the predominant factor. To summarize, there was no clear difference between students with and without LD.

Despite some slight individual differences, the results generally aligned with Toste et al. (2019), who found that automation is achieved through repetition and by presenting sub-lexical units with many words as examples. In summary, these results confirm indications from other studies that basic reading skills can be improved through phonics instruction, repeated reading, and word decoding (Donegan & Wanzek, 2021; Scammacca et al., 2015). Moreover, the results are consistent with previous racetrack studies (Barwasser et al., 2021a; Barwasser et al., 2021b; Barwasser et al., 2021c; Sperling et al., 2019), even though sub-lexical training was administered before the racetrack activity in the current intervention. Further, the implemented motivational boosters could also have had an effect, since motivation and learning are strongly linked (Scanlon et al., 2017). Even though some children did not enjoy partner work, all participants clearly profited from the intervention, and thus, results might also align with those of peer-tutoring studies (e.g., Bowman-Perrott et al., 2013; Moeyaert et al., 2021).

Limitations and Implications

In addition to the significant results, there are some limitations and implications to mention. The first limitation is that this is a single case study with only a few subjects, which limits the validity of the results. However, with the help of this design, it was possible to track individual progress effectively, as well as specifically identify which students an intervention seemed to benefit

and which students experienced problems. However, a future goal might be to evaluate the intervention in the context of single case studies on a greater number of children with different characteristics to obtain a clearer picture. With this enhanced clarity, a group design would enable researchers to make more generalized statements about the effectiveness overall. In this context, it would also be possible to compare the intervention with another reading intervention to see which works better. A larger study with more subjects could then compare the intervention with another intervention.

Overall, the goal would be to generalize the results to a larger group of students. Barwasser et al. (2021d) have already shown that a racetrack intervention is possible in an inclusive classroom This intervention could be repeated by including the sub-lexical units. Accordingly, it would be possible to optimize an intervention better than with a large group design. In addition, in the present study, there were already higher values in the baseline, and therefore, the upper limit in the B phase was, in principle, reached more quickly. However, 20 words were removed from a total pool of 70 for each measurement and the students did not practice and see the same 20 words each time. Thus, the increase from the A to the B phase is even more significant. Moreover, regarding the assessment, to avoid the possibility that the students improved because the assessment occurred immediately after the intervention, we recommend that future studies measure the dependent variable before the intervention sessions. However, experience shows that there is a risk of demotivation and frustration if students are measured beforehand with this intervention. Further, the intervention sessions did not provide any training in transfer words, and students still improved at these words. Nevertheless, a direct training effect on the dependent variables cannot be 100% excluded and should be considered.

Further, due to the multi-component intervention, it is unknown what effects the direct sub-lexical instruction had, what effects the racetracks alone had, or what effects the self-graphing had. However, as the method is easy to implement and it seems to work in combination, it would not necessarily be the goal to separate and evaluate the components. It is also not possible to clearly identify which characteristics were responsible for the response to the intervention and whether the characteristics played any role at all. Another possible limitation is that the students might have trained incorrectly on the words during the racetrack game, despite monitoring. However, because all students improved, one can assume that they generally did not train the words incorrectly. Nevertheless, as we did not focus on which words the students read correctly in a one-second rhythm, students may have experienced difficulties with specific words. Future research must consider this factor.

Some students had difficulty with transfers. In this case, we extended the time of the sessions to increase the intensity and repetition. According to

Richards-Tutor et al. (2016), research on the effects of literacy interventions on students with learning challenges or at risk for them is rare. Consequently, further research must involve students who experience severe challenges, especially with German L2. Although it is not possible to make a clear statement about mediators, when designing an intervention, the following factors require consideration: that the material is not too difficult, that some children prefer to work alone, and that vocabulary, academically engaged behavior, and especially motivation can be influencing factors in the effectiveness of an intervention. However, the only slight clues from the present study are behavioral problems associated with less proficient performance and LD in the context of transfer knowledge. Generally, the students had several at-risk components for learning difficulties, and we did not know which component might have influenced the intervention's success. However, against the background of inclusive education and the Convention on the Rights of the Child (UNICEF, 1989), this intervention could give hints of what a successful intervention can look like when applied to different kinds of students simultaneously. Nevertheless, future work must examine the different characteristics and the degree of influence on learning activities.

Another important point is that we did not collect follow-up data, which is important for determining whether results are long-lasting. Moreover, we did not control for time as an influential factor if follow-up data remained stable. However, collecting follow-up data was not possible due to the COVID-19 school closure. Nevertheless, future studies should collect follow-up data.

Conclusion

Overall, the combined intervention, which is easy to implement, was able to elicit improvements at both the sight word level and the transfer level. Nevertheless, the study provides important insights for integrating a simple reading exercise in a manner that may facilitate transfer effects on reading competence. The sample also consisted of very low-proficiency readers who additionally had LD, problem behavior, and/or German as L2. All children showed improvement, demonstrating an important step in reading research in this area. It is critical to consider inclusion, the creation of equal conditions for all students regardless of their characteristics, strengths, and weaknesses, and the fact that weak primary school readers face a higher risk of school dropout in secondary school (Brasseur-Hock et al., 2011). The intervention works for children with different language backgrounds and different personal characteristics.

References

Afracan, K., Wilkerson, K.L., & Ruppar, A.L. (2018). Multicomponent reading interventions for students with intellectual disability. *Remedial and Special Education, 39*(4), 229–242. https://doi.org/10.1177/0741932517702444

Albers, C. A., & Hoffman, A. (2012). Using flashcard drill methods and self-graphing procedures to improve the reading performance of English language learners. *Journal of Applied School Psychology, 28*, 367–388.

Ardoin, S. P., Morena, L. S., Binder, K. S., & Foster, T. E. (2013). Examining the impact of feedback and repeated readings on oral reading fluency: Let's not forget prosody. *School Psychology Quarterly*, *28*(4), 391–404. https://doi.org/10.1037/spq0000027

Barwasser, A., Hertel, S., & Grünke, M. (2021a). The effect of a (sub)lexical intervention using peer racetracks on the reading performance of low-literacy students with behavior problems with and without learning disabilities. *Learning Disabilities. A Contemporary Journal, 19*(2), 143-159.

Barwasser, A., Dufrene, B., Schmitz, S., Nobel, K., Spilles, M., & Grünke, M. (2021b). Effects of a reading racetrack intervention and additional self-graphing on reading and behavior of struggling readers with emotional behavioral disorders and learning disabilities. *European Journal of Special Needs Education, 38*(4), 511–527. *https://doi.org/10.1080/08856257.2022.2127080*

Barwasser, A., Nobel, K., & Grünke, M. (2021c). Peer-reading racetracks on word reading of low-achieving graduating students with learning disabilities and behavioral problems. *British Journal of Special Education.*

Barwasser, A., Urton, K., Grünke, M., Sperling, M., & Coker, D. Jr. (2021d). Fostering word fluency of struggling third graders from Germany through motivational peer-tutorial reading racetracks. *Reading & Writing*, https://doi.org/10.1007/s11145-021-10172-3

Barwasser,A., Hisgen, S., & Grünke, M. (2022). The effects of peer racetracks on the reading performance of German L2 Students with and without behavioral difficulties. *International Journal for Research in Learning Disabilities, 5*(2), 70-85. https://doi.org/10.28987/ijrld.5.2.70

Becker-Mrotzek, M., Hentschel, B., Hippmann, K., & Linnemann, M. (2012, June). *Sprachfoerderung in deutschen Schulen – die Sicht der Lehrerinnen und Lehrer. Ergebnisse einer Umfrage unter Lehrerinnen und Lehrern [Language support in German schools - the teachers' view. Results of a survey among teachers].*

Bowers, P. G. (1993). Text reading and rereading: Determinants of fluency beyond word recognition. *Journal of reading behavior*, *25*(2), 133–153.

Bowers, P. G., Sunseth, K., & Golden, J. (1999). The route between rapid naming and reading progress. *Scientific Studies of Reading*, *3*(1), 31–53.

Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., and Parker, R. (2013). Academic benefits of peer tutoring: a meta-analytic review of single-case research. Sch. Psychol. Rev. 42, 39–55. https://doi.org/ 10.1080/02796015. 2013.12087490

Brasseur-Hock, I. F., Hock, M. F., Kieffer, M. J., Biancarosa, G., & Deshler, D. D. (2011). Adolescent struggling readers in urban schools: Results of a latent class analysis. *Learning and Individual Differences*, *21*(4), 438–452. https://doi.org/10.1016/j.lindif.2011.01.008

Briesch, A. M., Chafouleas, S. M., Neugebauer, S. R., & Riley-Tillman, T. C. (2013). Assessing influences on intervention implementation: Revision of the Usage Rating Profile-Intervention. *Journal of School Psychology, 51*(1), 81–96. https://doi.org/10.1016/j.jsp.2012.08.006

Byiers, B. J., Reichle, J., and Symons, F. J. (2012). Single subject experimental design for evidence-based practice. *Am. J. Speech Lang. Pathol. 21*, 397–414. doi: 10.1044/1058-0360(2012/11-0036)

Cole, M. W. (2014). Speaking to read: Meta-analysis of peer-mediated learning for English language learners. *Journal of Literacy Research, 46*(3), 358–382.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. **Psychological Review, 100**(4), 589–608. https://doi.org/10.1037/0033-295X.100.4.589

Cop, U., Drieghe, D., & Duyck, W. (2015). Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. *PloS One*, *10*(8), 1–38.

De Jong, C. G., Licht, R., Sergeant, J. A., & Oosterlaan, J. (2012). RD, ADHD, and their comorbidity from a dual route perspective. *Child Neuropsychology*, *18*(5), 467–486.

Diamanti, V., Goulandris, N., Campbell, R., & Protopapas, A. (2018). Dyslexia profiles across orthographies differing in transparency: An evaluation of theoretical predictions contrasting English and Greek. *Scientific Studies of Reading*, *22*(1), 55–69.

Donegan, R. E., & Wanzek, J. (2021). Effects of reading interventions implemented for upper elementary struggling readers: A look at recent research. *Reading and Writing*. https://doi.org/10.1007/s11145-021-10123-y

Dufrene, B. A., Reisener, C. D., Olmi, D. J., Zoder-Martell, K., McNutt, M. R., & Horn, D. R. (2010). Peer tutoring for reading fluency as a feasible and effective alternative in response to intervention systems. *Journal of Behavioral Education, 19*(3), 239–256.

Dunn, M. E., Shelnut, J., Ryan, J. B., & Katsiyannis, A. (2017). A systematic review of peermediated interventions on the academic achievement of students with emotional/behavioral disorders. *Education and Treatment of Children, 40*(4), 497–524.

Ehri, L. C. (2005). Development of sight word reading: Phases and findings. In M. J. Snowling & C. Hulme (Eds.), *Blackwell handbooks of developmental psychology. The science of reading: A handbook* (pp. 135–154). Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch8

Gangl, M., Moll, K., Jones, M., Banfi, C., Schulte-Körne, G., & Landerl, K. (2018). Lexical reading in dysfluent readers of German. **Scientific Studies of Reading, 22(1)**, 24–40. https://doi.org/10.1080/10888438.2017.1339709

Gorsuch, G., & Taguchi, E. (2008). Repeated reading for developing reading fluency and reading: The case of EFL learners in Vietnam. *System, 36*(2), 253–278.

Goswami, U., Ziegler, J., & Richardson, U. (2006). The effects of spelling consistency on phonological awareness: A comparison of English and German. *Journal of Experimental Child Psychology*, *92*(4), 345–365. https://doi.org/10.1016/j.jecp.2005.06.002

Grünke, M., & Cavendish, W. (2016). Learning disabilities around the globe: Making sense of the heterogeneity of the different viewpoints. *Learning Disabilities: A Contemporary Journal*, *14*, 1–8.

Guzman, G., Goldberg, T. S., & Swanson, H. L. (2018). A meta-analysis of self-monitoring on reading performance of K–12 students. *School Psychology Quarterly, 33*(1), 160–168.

Harju-Luukkainen, H., McElvany, N., & Stang, J. (2020). Monitoring student achievement in the 21st century Europena policy perspectives and assessment strategies. Springer Nature Switzerland AG.

Huemer, S., Landerl, K., Aro, M., & Lyytinen, H. (2008). Training reading fluency among poor readers of German: Many ways to the goal. *Annals of Dyslexia, 58*(2), 115–137.

Kazdin, A. E. (2010). *Single-case research designs: Methods for clinical applied settings.*

Knoepke, J., Richter, T., Isberner, M. B., Naumann, J., & Neeb, Y. (2014). Phonological recoding, orthographic decoding, and comprehension skills during reading acquisition. *Zeitschrift Für Erziehungswissenschaft, 17,* 447–471. https://doi.org/10.1007/s11618-014-0560-z

Lämsä, J., Hämäläinen, R., Aro, M., Koskimaa, R., & Äyrämö, S.-M. (2018). Games for enhancing basic reading and maths skills: A systematic review of educational game design in supporting learning by people with learning disabilities. **British Journal of Educational Technology**, *49*(4), 596–607. https://doi.org/10.1111/bjet.12639

Landerl, K., Freudenthaler, H. H., Heene, M., De Jong, P. F., Desrochers, A., Manolitsis, G., ... & Georgiou, G. K. (2019). Phonological awareness and rapid automatized naming as longitudinal predictors of reading in five alphabetic orthographies with varying degrees of consistency. *Scientific Studies of Reading*, *23*(3), 220–234.

Lee, J., & Yoon, S. Y. (2015). The effects of repeated reading on reading fluency for students with reading disabilities: A meta-analysis. *Journal of Learning Disabilities*, *50*(2), 213–224. https://doi.org/10.1177/0022219415605194

Lervåg, A., & Aukrust, V. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension between first and second language learners. *Journal of Child Psychology and Psychiatry*, *51*, 612–620.

Ma, H. H. (2006). An alternative method for quantitative synthesis of single subject researches: Percentage of data points exceeding the median. *Behavior Modification*, *30*(5), 598–617. https://doi.org/10.1177/0145445504272974

Manolov, R., & Moeyaert, M. (2017). Recommendations for choosing single-case data analytical techniques. *Behavior Therapy*, *48*(1), 97–114. https://doi.org/10.1016/j.beth.2016.04.008

Mayer, A. (2018). *Blitzschnelle Worterkennung [Rapid word recognition]* (3rd ed.). Borgmann Media.

Mayringer. H. & Wimmer, H. (2003). Salzburger Lesescreening (SLS 1-4) [*Salzburg Reading Screening*]. Bern: Verlag Hans Huber.

McElvany, N., Lorenz, R., Frey, A., Goldhammer, F., Schilcher, A., & Stubbe, T.C. (2021). IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre [*Reading literacy of primary school children in international comparison and trend over 20 years*]. Waxmann.

McKenna, J. W., & Bettini, E. (2018). Improving reading fluency skills for secondary students with emotional and behavioral disorders. *Beyond Behavior, 27*(2), 74–81.

Melby-Lervåg, M. & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological Bulletin, 140*(2), 409–433.

Moll, K., & Landerl, K. (2010). *SLRT-II: Lese-und Rechtschreibtest; Weiterentwicklung des Salzburger Lese-und Rechtschreibtests (SLRT) [Reading and spelling test; further development of the Salzburg reading and spelling test]*. Huber.

Moeyaert, M., Klingbeil, D. A., Rodabaugh, E., and Turan, M. (2021). Three level meta-analysis of single-case data regarding the effects of peer tutoring on academic and social-behavioral outcomes for at-risk students and students with disabilities. *Remedial Spec. Educ. 42:074193251985507*. doi: 10.1177/ 0741932519855079

Musti-Rao, S., Hawkins, R. O., & Barkley, E. A. (2009). Effects of repeated readings on the oral reading fluency of urban fourth-grade students. *Preventing School Failure, 54*(1), 12–23.

Nelson, J., & Harwood, H. (2010). Learning disabilities and anxiety: A meta-analysis. *Journal of Learning Disabilities*, *44*, 3–17.

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single case research: A review of nine non-overlap techniques. *Behavior Modification*, *35*(4), 302–322. https://doi.org/10.1177/0145445511399147

Peterson, R. L., Boada, R., McGrath, L. M., Willcutt, E. G., Olson, R. K., & Pennington, B. F. (2017). Cognitive prediction of reading, math, and attention: Shared and unique influences. *Journal of Learning Disabilities*, *50*(4), 408–421. https://doi.org/10.1177/0022219415618500

Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher, 58*(6), 510–519.

Prowse, P. T., & Nagel, T. (2015). A meta-evaluation: The role of treatment fidelity within psychosocial interventions during the last decade. *Journal of Psychiatry*, *18*(2), 1–7.

Rashotte, C. A., & Torgesen, J. K. (1985) Repeated reading and reading fluency in learning disabled children. *Reading Research Quarterly*, *20*(2), 180–188.

Reid, R., Ortiz Lienemann, T., & Hagaman, J. L. (2013). *Strategy instruction for students with learning disabilities* (2nd ed.). The Guilford Press.

Richards-Tutor, C., Baker, D. L., Gersten, R., Baker, S. K., & Smith, J. M. (2016). The effectiveness of reading interventions for English learners: A research synthesis. *Exceptional Children, 82*, 144 –149.

Roberts, G. J., Cho, E., Garwood, J. D., Goble, G. H., Robertson, T., & Hodges, A. (2020). Reading interventions for students with reading and behavioral difficulties: A meta-analysis and evaluation of co-occurring difficulties. *Educational Psychology Review*, *32*(1), 17–47. https://doi.org/10.1007/s10648-019-09485-1

Roberts, G. J., Solis, M., Ciullo, S., McKenna, J. W., & Vaughn, S. (2015). Reading interventions with behavioral and social skill outcomes: a synthesis of research. *Behavior modification*, *39*(1), 8–42. https://doi.org/10.1177/0145445514561318

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities*, *48*(4), 369–390. https://doi.org/10.1177/0022219413504995

Scanlon, D. M., Anderson, K. L., & Sweeney, J. M. (2017). *Early intervention for reading difficulties: The interactive strategies approach*. Guilford.

Solis, M. M., Ciullo, S., Vaughn, S., Pyle, N., Hassaram, B., & Leroux, A. (2012). Reading comprehension interventions for middle school students with learning disabilities: A synthesis of 30 years of research. *Journal of Learning Disabilities*, *45*, 327–340.

Sperling, M., Barwasser, A., & Grünke, M. (2019). The effects of a reading racetrack intervention on the sight word fluency of learning-disabled elementary school students with German as second language. *Insights into Learning Disabilities, 16*(1), 79–90.

Statistisches Bundesamt (2022). Ergebnisse des Mikrozensus 2020. Fachserie 1 Reihe 2.2, S. 49, eigene Berechnung; Statistisches Bundesamt (2019): Was ist der Mikrozensus?

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.

Toste, J., Capin, P., Williams, K., Cho, E., & Vaughn, S. (2019). Replication of an experimental study investigating the efficacy of a multisyllabic word reading intervention with and without motivational beliefs training for struggling readers. *Journal of Learning Disabilities*, *52*(1), 45–58. https://doi.org/10.1177/0022219418775114

UNICEF UK. (1989). *The United Nations convention on the rights of the child.* https://downloads.unicef.org.uk/wp-content/uploads/2010/05/UNCRC_PRESS200910web.pdf?_ga=2.78590034.795419542.1582474737-1972578648.1582474737

Volpe, R. J., Casale, G., Mohiyeddini, C., Grosche, M., Hennemann, T., Briesch, A. M., & Daniels, B. (2018). A universal behavioral screener linked to personalized classroom interventions. Psychometric characteristics in a large sample off German schoolchildren. *Journal of School Psychology, 66*, 25–40. https://doi.org/10.1016/j.jsp.2017.11.003

Weiß, R. H. (2006). *Grundintelligenztest Skala 2 - Revision (CFT 20-R) mit Wortschatztest und Zahlenfolgentest. [Basic Intelligence Test Scale 2 - Revision (CFT 20-R) with vocabulary test and number sequence test.]* Hogrefe.

Wu, S., Stratton, K. K., & Gadke, D. L. (2020). Maximizing repeated readings: The effects of a multicomponent reading fluency intervention for children with reading difficulties. *Contemporary School Psychology, 24*(2), 217–227.

Zentall, S. S. (2014). *Students with mild exceptionalities: Characteristics and applications*. Sage. https://doi.org/10.4135/9781506374604

Author's Note

Correspondence concerning this article should be addressed to Anne Barwasser, Ph.D., Department of Special Education & Rehabilitation, University of Cologne, Klosterstr. 79b, Cologne, Northrhine-Westfalia, 50931, Germany, E-mail: anne.barwasser@uni-koeln.de.