A (Sub)-Lexical Patterns Training With Racetracks on Trained and Untrained Words in Low-Literacy German Students With Behavior Problems With and Without Learning Disabilities

Anne Barwasser*
Sophia Hertel
Matthias Grünke
University Cologne, Germany

The ability to read adequately is one of the most important skills for students to achieve during their school career. Unfortunately, a large number of children do not have adequate reading skills and thus may face problems in many respects throughout their lives. Children with learning disabilities (LD) and behavioral problems are particularly prone to fall into this category and may additionally experience a general lack of motivation to learn. In the context of improving reading skills, fostering lexical and sublexical reading has proven successful. This single-case study (N = 5) investigated the effect of using a motivating intervention consisting of a (sub)lexical patterns training in combination with reading racetracks for the automation of common German sublexical patterns to improve the reading of trained and untrained words. The intervention, which was carried out three times a week for 20 minutes each over a period of five weeks, showed promising results, also in the follow-up measurements, including medium to strong effects on training words and small to strong effects on transfer words. These findings offer preliminary evidence of how to combine reading racetracks to create transfer effects in reading for low-achieving primary school students with severe learning and behavioral issues.

Keywords: **Peer Tutoring, Reading Racetracks; Lexical and Sub-lexical Reading; Learning Disabilities; Behavioral Problems**

Introduction

The Significance of Reading Fluency and Sight Word Recognition

The ability to decode written and printed symbols is a key skills that has a decisive influence on almost all areas of life (Grigoryan, 2020; Macdonald et al., 2016). Regrettably, about 10-15% percent of those who have received at least basic education worldwide have major difficulties in understanding written texts (Dyslexia Action, 2017). In Germany, studies have found an upward trend of fourth graders with poor reading competency accompanied by declining reading motivation (Bos et al., 2017). In particular, many students with learning disabilities (LD) face challenges with regard to acquiring reading competency (Lerner & Johns, 2011; Solis et al., 2012), which can be partly linked to lower memory capacity, especially regarding

*Please send correspondence to: Anne Barwasser, MA, Department of Special Education & Rehabilitation, University of Cologne, Klosterstr. 79b, 50931 Cologne, Germany, Email: anne.barwasser@uni-koeln.de.

word reading (Geary et al., 2020). In addition, there is a relationship between behavioral problems and reading difficulties. That is, students who demonstrate behavioral challenges spend less time reading (Vaughn et al., 2002), and many perform at least a whole school year behind their peers (Oakes et al., 2010).

When reading fluently, an experienced reader retrieves words automatically from the mental lexicon (Morris & Perney, 2018; Young et al., 2020). The ability to read fluently is indispensable for reading comprehension and, thus, for overall reading proficiency (Ehri, 2005). According to the dual-route theory (DRT; Coltheart, 2005), the reading process happens through a lexical and a nonlexical route (sublexical). In the nonlexical route, graphemes are recorded and converted into phonemes (phonological recording), whereas in the lexical route, written words are mapped directly on to mental representations of word forms (orthographical decoding) (Coltheart et al., 2005). If students can process frequently occurring orthographic patterns of the German language, for example, this can guide direct word recognition of a large number of words (Mayer, 2018).

Experienced readers have built up these so-called "sight words," which enables them to recognize a word from the mental lexicon within one second of its appearance (Ehri, 2005). Knoepke et al. (2014) demonstrated that readers of German (a transparent language; that is, there is a one-to-one relationship between meaning and form), orthographic decoding is a greater predictor of reading skills than phonological recoding (see also Tressoldi et al., 2007). According to Ehri (2014), being able to match sounds to orthographic patterns is critical for word recognition. Thus, nonlexical decoding skills also play a major role with respect to reading fluency (Harn et al., 2008). Therefore, combining both in an intervention seems to be a promising way to foster reading performance.

(Sub)Lexical Training and Repeated Reading

Reading instruction should start at the word level (Ehri, 2014). According to a meta-analysis by Scammacca et al. (2007), any student can profit from a word-level intervention. Further, word-recognition interventions have shown positive results (Martin-Chang & Levy, 2005; Scammacca et al., 2015). Going a step further, apart from whole-word training, Marinus et al. (2012) found that explicit training of sublexical patterns led to superior short- and long-term improvement in the rapid naming of trained and untrained patterns for primary students with poor reading skills. To become fluent readers, people need to read words or word patterns many times (Grabe, 2010); hence Repeated Reading (RR) has been found to be a promising tool for improving reading skills (Kostewicz et al., 2016; Lee & Yoon, 2017; Rasinski et al., 2016). The repetition of words and word patterns allows readers to automatically retrieve them from their mental lexicon (Zavala & Cuevas, 2019), thereby relieving the demands on working memory.

Reading Racetracks as a Realization of RR and (Sub)Lexical Training

Research has indicated that Reading Racetracks (RT) is an effective and enjoyable way to improve reading fluency through RR (Crowley et al., 2013). As a game-based instructional method, RT uses a game board consisting of cells that contain items such as sight words. Numerous studies have shown the effectiveness of

racetracks for both first-language (L1) and second-language (L2) students (Grünke, 2019; Grünke & Barwasser, 2019; Sperling et al., 2019); in addition, when combined with peer tutoring (PT) and motivational components, RT has been found effective for a wide range of students with and without LD (e.g., Barwasser et al., 2021a, 2021b). However, existing studies have not focused on possible transfer effects – solely on trained items.

Peer Tutoring as Inclusion Tool

Given the increase in inclusive educational programming, peer tutoring (PT), an evidence-based intervention whereby students are trained to provide instruction to their peers (Bertin & Narcy-Combes, 2007), is becoming a popular intervention choice. For example, PT has been found to be successful for promoting learning in collaborative and inclusive environments (Alzahrani & Leko, 2018). Of particular interest in the current context, multiple reading strategies have been effectively taught in peer-assisted learning arrangements (Sáenz et al., 2007; Van Keer & Vanderlinde, 2010). Overall, students who were integrated into a reading intervention that included PT performed better than students who only received teacher-centered reading instruction (Mueller et al., 2015). To our knowledge, RT have been only combined with PT in the studies by Barwasser et al. (2021a, 2021b).

Incorporation of Self-Graphing

Motivation also plays an important role in reading (Marinak & Gambrell, 2008). Students who lack motivation to engage in activities such as reading try to avoid them (Guthrie, 2000; Marinak & Gambrell, 2008); hence motivation must be given serious consideration in efforts to improve reading skills. Motivation techniques in reading interventions are particularly important for students who experience high levels of frustration and, therefore, are at risk of failing (Sideridis, 2002).

One possible technique is self-graphing, which is a specific type of self-evaluation wherein students write down their own performances, thus creating a visual representation of their performance over time (Gunter et al., 2002). In a meta-analysis on the effects of formative evaluation, Fuchs and Fuchs (1986) noted a significant increase in the academic achievement of students with disabilities through systematic formative evaluation procedures, especially when the data were presented graphically (see also Sutherland & Snyder, 2007).

Research Aim

Given that the gap between skilled and less skilled readers is widening in Germany, and indeed throughout the world, and that children with behavioral problems and LD face special challenges in his respect, it is of great importance to develop an intervention of a motivating and automating character and that can be used for a wide variety of students. Combining the aforementioned methodological aspects into one intervention, as done by Barwasser et al. (2021a, 2021b), the current study focused on the effects of a combined racetrack intervention on trained and, as a unique selling point, untrained words, to determine if the intervention is a universal tool that additionally can lead to transfer effects. Moreover, the social validity of the intervention – that is, how it is received by students – was examined.

The following research questions underlie the study: (a) Did the combined (sub)lexical patterns racetrack intervention with PT and self-graphing lead to an increase in the acquisition of trained and untrained (transfer) words? (b) Were the results stable at two-month follow-up? (c) How was the intervention received by the students?

Methods

Participants and Setting

The study was conducted in an urban elementary school in North Rhine-Westphalia, Germany. Students from third and fourth grade with low proficiency in word reading were targeted. Prior to the start of the study, consent to participate in the study was obtained from students' legal guardians.

To select the final participants, a multi-step procedure was used as follows. (a) A German reading screening that focuses on word reading (SLRT II; Moll & Landerl, 2010) was administered to the whole classes. The SLRT II consists of two 1-minute word and pseudo-word reading tests. The SLRT II's correlation with other reading screenings lies between .69 and .92, and its reliability is between .90 and .98. Students who achieved a percentile (PR) of <15 on both tests were eligible for the study (see Table 1). (b) A German vocabulary test (WS/ZF-R; Weiß, 2006) and a screening to assess externalizing problem behavior (Integrated Teacher Report Form [ITRF]; Volpe et al., 2018) were also used. The WS/ZF-R examines areas of crystallized ability (skills acquired through prior knowledge and experience); it contains a total of 30 items including words from semantic fields as well as abstract terms. The relevant task is to select from a series of five words the one that has the same or similar meaning as a given word. The WS/ZF-R has "good" to "very good" reliability (.87). The ITRF-G (short form) is a German translation of the American screening Integrated Teacher Report Form; it is considered by Volpe et al. (2018) as a universal and instructionally relevant behavioral screening to crystallize externalizing behavior. It consists of a total of 16 items, with eight items each assigned to the areas of "problems in learning behavior" and "oppositional behavior;" the cutoff value for the total problem behavior is 13. The ITRF has a high internal consistency (total problem value: $\alpha = .91$). In addition, a teacher questionnaire was distributed to assess further characteristics of the children.

The final sample consisted of 8 students, divided into three groups, each with different baseline lengths. The age range of students was between 8 and 10 years old. All children showed low proficiency in word reading and German vocabulary. Also, all children exhibited general problem behavior according to the ITRF, and two were diagnosed with LD. In Germany, LD means immense difficulties in more than one school subject mostly accompanied by a rather low IQ (Hasselhorn & Gold, 2017). Due to the missing data (as a result of COVID-19 quarantine), this article focuses on the data of five students ($N = 5$). Reading pairs were formed based on teacher assessment and results of reading screening (that is the pairs of students had the same reading level).

Table 1. ***Demographic Characteristics of the Participants***

Name	Gender	Age	Reading W/PW PR	Vocabulary PR	ITRF GP	SEN	German L2	Ethnicity
Latifa	female	10	7-8/4	10	21	LD	Yes	Lingala
Ella	female	9	<1/<1	2	18	LD	Yes	Greek
Lou	female	8	4-5/<2	0	15	/	No	German
Allai	female	8	1-2/14	2	14	/	Yes	Arabic
Jim	male	8	11-15/9-11	1	20	/	Yes	Thai

Note. Words (W); Pseudo Words (PW); Percentile (PR, cutoff 15); Integrated Teacher Report Form (ITRF); Main Problem Value (GP); Special Educational Needs (SEN); Second Language (L2); Learning Disabilities (LD).

Design

A multiple-baseline design across participants was applied to be able to exclude alternative explanations for the effectiveness of the intervention (Kazdin, 2010). The children were randomly assigned to the three groups; due to the COVID-19 pandemic, Group 3 and 4 could not be supported together. Group 1 had a total of five baseline measurements, Group 2, four, and Group 3, six (Kratochwill et al., 2013). In total there were 20 planned measurement points for baseline and intervention together. Before the baseline started, screenings were performed within 1.5 weeks. The baseline phase was then started three times a week, followed by the intervention phase, which also started three times a week over a period of five weeks. Data were collected after each baseline and each intervention session, as well as two months after the end of the intervention (follow-up). The groups were brought, one after the other, to rooms outside the students' regular classroom. As test leaders and interventionists, six master's-level students of special needs education were employed. Four students supported the groups in pairs. The other two were responsible for the measurements to avoid bias.

Dependent Variables and Measurements

There were two dependent variables: (a) number of correctly read training words and (b) number of correctly read transfer words. A researcher-developed instrument was used, which consists of a PowerPoint presentation (PPT) in which a word to be read is shown on each slide. The slides are set to a 1-second rhythm (Ehri, 2005). After each baseline and each intervention session, the participants were assessed using this PPT. There were two pools of words, 70 training words and 70 transfer words, taken from "Rapid Word Recognition" (Mayer, 2018). Twenty words from each pool were used randomly in the PPT, which allowed the students to reach a maximum of 20 words for training words and 20 words for transfer words for each measurement. Care was taken to ensure that the 70 words were distributed equally among the measurement points and that the measurement points themselves did not differ significantly from each other with regard to word difficulty, as shown by a variance analysis using SPSS statistics. Afterwards the number of correctly read training and transfer words was noted.

Procedure

Baseline

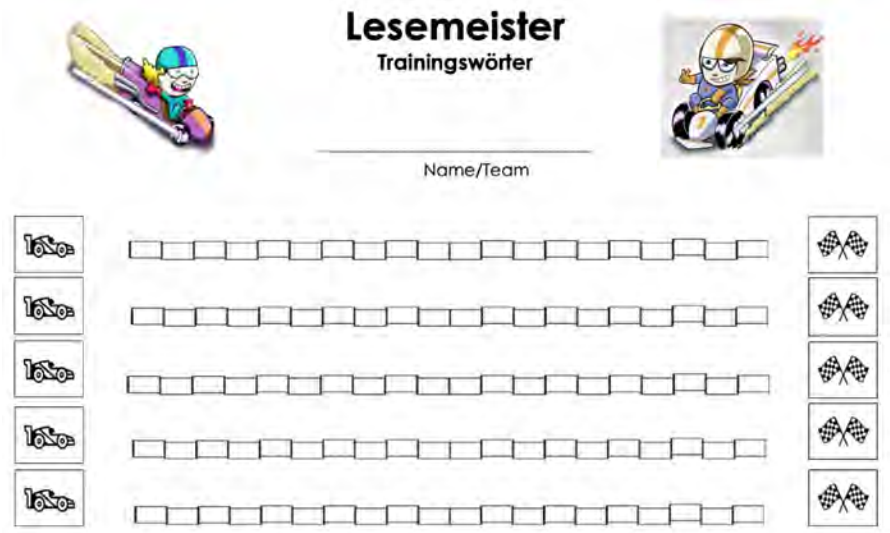
To engage the children in the baseline (Phase A), cognitive tasks were applied, consisting of logically continuing a certain sequence and crossing out symbols that do not fit into a row. These exercises were chosen since they do not foster reading, allowing the current state of the two dependent variables better to be estimated. Baseline was conducted in the same group as in the intervention. After 20 minutes, all participants were assessed independently with respect to the dependent variables.

Intervention

The intervention was made up of two stages. The first stage consisted of direct instruction of letter clusters and training words using large flashcards while the children sat in a semicircle in front of the interventionist for 10 minutes. For this stage, the materials of the German program *Blitzschnelle Worterkennung* (Rapid Word Recognition; Mayer, 2018) were used. This involved a pool of training words (70) and transfer words (70) as well as letter clusters. The letter clusters and matching training words were printed individually on 8.3 x 11.7-inch flashcards; the letter cluster was marked in green for the training words. The letter clusters were slowly presented, one by one, and then the appropriate training words were presented. The interventionist read all the words aloud, and the children were also asked to try to read the words and clusters.

The second stage consisted of the RT procedure where the pairs of students previously formed played a game for 10 minutes. A 11.7 × 16.5-inch game board with 14 empty cells was designed. In addition, the training words were printed individually on small flashcards that were distributed across the racetrack. The training words (always 14 out of the pool of 70) are placed upside down on each cell of the racetrack field, and the children take turns throwing a die. Child 1 throws a die and moves the game figure according to the number of eyes on the die and lifts the corresponding flascard. The child tries to read the word and Child 2 corrects if possible. If the word was not read correctly, Child 2 reads the word and Child 1 repeats it. Now, both children think about which letter cluster is recognizable in the word. If the word is read correctly, students move right on while child 2 repeats the same process. The children collect the correctly read flashcards. Not correctly read words remain on the board. If both children cannot read the word correctly, the interventionists will help.

After each intervention session, all participants are assessed independently with respect to the dependent variables. To increase motivation, the children are asked to enter the number of correctly read training and transfer words in two self-graphing sheets after each measurement. Each sheet consists of several rows, one below the other (number of sessions), each with 20 boxes, representing the maximum number possible of correctly read words (see Figure 1).



Treatment Fidelity and Social Validity

Prior to the study, the interventionists and test leader were intensively trained in how to conduct the study. For treatment fidelity, a documentation sheet was designed to be filled out by the interventionists after each session and, additionally, by an external person for one third of the intervention time. The questionnaire contained a checklist consisting of the following items: (a) environment/external circumstances, (b) planning, (c) material, (d) course of the intervention, (e) diagnostics, and (f) feedback and dealing with students. Items were ranked on a 5-point Likert scale from 0 (not applicable at all) to 4 (completely applicable). At the end of the checklist there is room for adding comments and/or remarks. The interrater reliability was 100%, both between the interventionists and between the interventionists and the external person across all groups.

To determine social validity, a questionnaire was designed consisting of 12 items to be rated by the participants on a 5-point Likert scale ranging from 0 = not true at all to 4 = completely true. *Item 1*: The racetrack game helped me to read words correctly; *Item 2*: I think that the support also helps other students with reading difficulties; *Item 3*: I have understood the purpose of the intervention well; *Item 4*: I learned a lot during the program; *Item 5*: I enjoyed coming to the program; *Item 6*: I enjoyed the support; *Item 7*: I would participate in the program again. *Item 8*: The letter clusters groups helped me to read better; *Item 9*: Drawing in the arcs was fun. *Item 10*: I would like to do something like that more often; *Item 11*: The words were difficult. *Item 12*: I enjoyed playing in pairs.

Results

Descriptive Analysis

Training Words

The statistical program R was used throughout. Additionally, the mean baseline difference (MBD) was calculated by hand. Regarding the dependent variable training words, there was a visual increase from Phase A to Phase B for each participant; same pattern is reflected in the descriptive data (see Table 2). Latifa, Allai, and Jim reached the maximum value of 20 correctly read training words during Phase B, followed by Ella and Lou with 18, with Lou showing the greatest increase. Moreover, the follow-up data (Phase E) were stable across all participants. The mean values showed a slight increase compared to Phase B. The MBD from Phase A to B was as follows: Jim: 21,34 %, Allai: 29.30 %, and Latifa: 31.00%, with a greater increase for Ella: 91,41% and Lou: 200%. With respect to mastery, Latifa needed six sessions, Allai four sessions, and Jim three sessions. Ella reached her maximum value in Session 7 and Lou in Session 12.

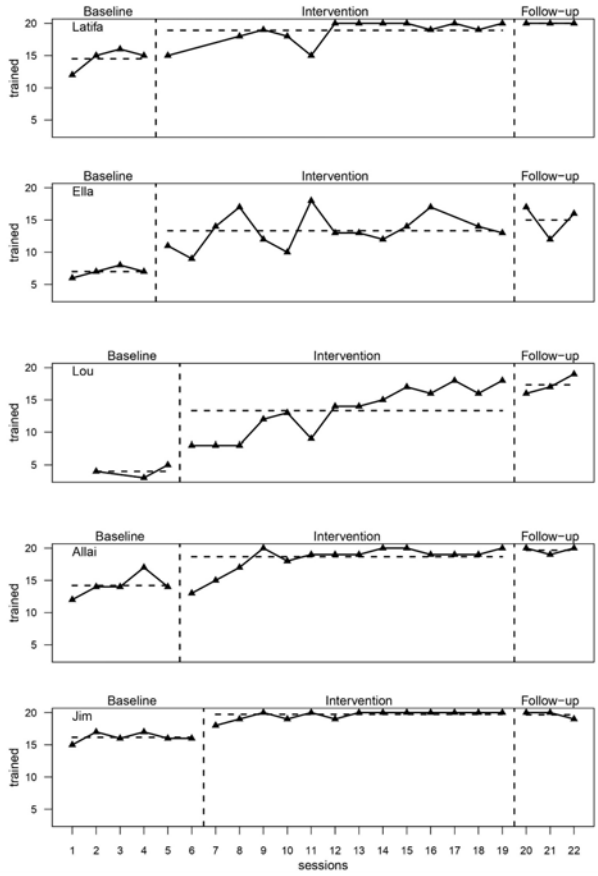


Table 2. ***Descriptive Data for Each Participant in A, B, and E Phase Training Words***

	N (A)	*N* (B)	*N* (E)	MA (*SD*)	MB (*SD*)	MBD	ME(*SD*)	Max B
Latifa	4	15	3	14.50(21.73)	19.00(1.80)	31,00%	20.00(0.00)	20.00
Ella	4	15	3	7.00(0.82)	13.38(2.72)	91,14%	14.67(2.52)	18.00
Lou	5	15	3	4.00(1.00)	12.00(3.84)	200,00%	17.33(1.53)	18.00
Allai	5	14	3	14.20(1.79)	18.36(2.06)	29,30%	19.67(0.58)	20.00
Jim	6	13	3	16.17(0.75)	19.62(0.65)	21,34%	19.67(0.58)	20.00

Note. Measurements (*N*); Baseline (A); Intervention (B); Follow-Up (E); Mean (*M*); Standard Deviation (*SD*); Maximum (Max); Mean Baseline Difference (MBD).

Transfer Words

For the transfer words, there was a clear difference from the training words since visually all children increased more slowly in Phase B than with the training words (see Table 3). The descriptive data show that Latifa and Jim reached a maximum value (20 in Phase B, followed by Allai with 19, Ella with 13, and Lou with 12, again with the largest increase.

For the follow-up measurements, the data were relatively stable as for the training words. Overall, the mean values showed a slight increase from Phase B to Phase E. Regarding the MBD, Latifa and Jim showed the weakest increases of 25.81% and 26.25 % from Phase A to Phase B, respectively, followed by Ella (44.60%) and Allai (48.10%). Lou, again, had the greatest improvement, with 204.35 %. Regarding mastery, Latifa needed 12 sessions and Jim, 11. Ella's maximum value was reached in Session 4, Allai's in Session 9, and Lou's in Sessions 13.

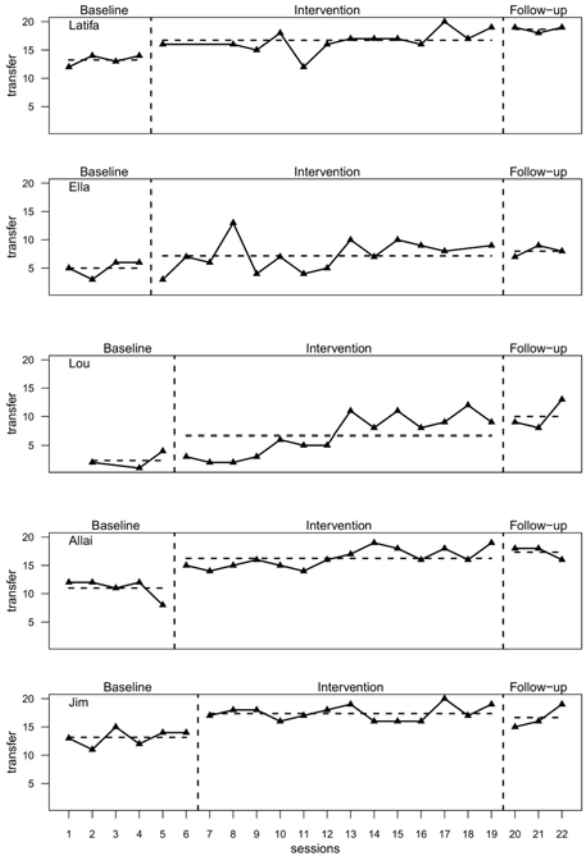


Table 3. ***Descriptive Data for Each Participant in A, B, and E Phase Transfer Words***

	N (A)	*N* (B)	*N* (E)	MA *(SD)*	MB *(SD)*	MBD	ME*(SD)*	Max B
Latifa	4	15	3	13.25(0.96)	16.67(2.02)	25,81%	18.67(0.58)	20.00
Ella	4	14	3	5.00(1.41)	7.23(2.89)	44,60%	8.00(1.00)	13.00
Lou	5	13	3	2.33(1.29)	7.00(3.44)	204,35%	10.00(2.65)	12.00
Allai	5	14	3	11.00(2.83)	16.29(1.68)	48,10%	17.33(1.15)	19.00
Jim	6	13	3	13.83(2.14)	17.46(1.33)	26,25%	16.67(2.08)	20.00

Note. Measurements (*N*); Baseline (A); Intervention (B); Follow-Up (E), Mean (*M*); Standard Deviation (*SD*); Maximum (Max); Mean Baseline Difference (MBD).

Overlap Indices

Training Words

Overlap indices were used for further and more in-depth analysis. The Non-Overlap of All Pairs (NAP; Parker et al., 2011) and Percentage of All Non-Overlapping Data (PAND) Parker et al. (2007) were chosen as well as Tau-U due to possible A-Phase trends.

Regarding training words, Latifa (92.00; $p<.01$) and Allai (92.00, $p<.01$) achieved a medium effect for NAP while Ella (100.00, $p<.001$), Jim (100.00, $p<.001$), and Lou (100.00) displayed a strong effect. Latifa (76.47) and Allai (89.47) showed moderate effects for PAND while Ella (100.00), Jim (100.00), and Lou (100.00) displayed strong effects. The Tau-U showed a moderate effect for Latifa (0.43, $p<.01$), Ella 0.46, $p<.01$), and Allai (0.57, $p<.001$) and a large change for Jim (0.64, $p<.001$) and Lou (0.68, $p<.001$).

Transfer Words

For transfer words, the NAP value was 77.00 ($p<.05$) for Ella and 88.00 for Lou ($p<.05$), signaling a moderate change. Latifa (93.00, $p<.01$), Jim 100.00, $p<.001$), and Allai (100.00, $p<.001$) showed a high effect strength. PAND values resulted in a small effect for Ella (66.71) and a moderate effect for Lou (79.47) and Latifa (82.35). A highly effective treatment was reflected by Allai (100.00) and Jim (100.00). Taking into account a possible A-Phase trend, Tau-U values showed a moderate effect for Ella (0.43, *<.05*), Latifa (0.43, $p<.05$), Jim (0.47, $p<.05$), and Lou (0.54, $p<.001$) and a large change for Allai (0.76, $p<.001$).

Social Validity

The results of the social validity questionnaire indicated an overall very positive student attitude towards the intervention, with Items 5-8 (*Item 5*: I enjoyed coming to the program; *Item 6*: I enjoyed the support; *Item 7*: I would participate in the program again. *Item 8*: The letter clusters helped me to read better) rated highest with an overall score of 4 (completely correct). Further, the students rated the words as not too difficult overall. Finally, the children's self-written comments show that all felt very well about the intervention, had fun, and would like to participate again (Latifa: "I would like to participate again in any case. It was a lot of fun."; Ella: "*It was fun and I would like to participate again.*"; Lou: "*I had much fun. Thank you.*"; Allai: "*I miss you already. Thanks for helping us. Your game was a lot of fun.*"; Jim: "*I had much fun playing the racetrack game.*").

Discussion

Main Findings

The aim of this study was to facilitate the reading of training and transfer words of children with severe reading difficulties with behavioral problems with and without LD using frequently occurring German letter clusters. Since it is of particular importance to create an intervention that can be used across a wide range of students, the present results are promising. Automatic decoding of words plays an immensely

important role in reading (Burns, 2007; Knoepke et al., 2014), and given that the gap between less proficient and stronger readers is widening (Bos et al., 2017), it is necessary to find effective interventions. Above all, reading acquisition in primary school is critical (Musti-Rao et al., 2015) as a preventive measure. Overall, the variable training words showed moderate to strong effects, and also proved to be statistically significant. A main focus here was on Tau-U, which takes into account an A-Phase trend that was present purely visually in some participants' performance. Three of the children reached the maximum value of 20 (mastery) in Phase B, closely followed by Ella and Lou with a value of 18. Unfortunately, Latifa, Allai, and Jim started with quite high values in Phase A; as a result, the maximum value of 20 in Phase B was quickly reached, and a ceiling effect was recognized for all three.

The follow-up data show stable values for all children. However, it should be noted that the pool of training words consisted of 70 words of which 20 were randomly selected for the measurement each time. Presumably, stronger effects would be achieved, if the number of training words were reduced. Nevertheless, the pool was large enough to avoid a single effect from the measurements.

For the variable transfer words, which is supposed to show whether the students could read untrained words through the clusters, as expected, lower effects were found, since it is more difficult to apply knowledge to unknown content than to retrieve known information. Here, the children started in Phase A with significantly lower values, whereby Latifa, Allai, and Jim again displayed the highest values. The increase in Phase B was visibly less steep than for the training words, but the Tau-U results showed moderate to strong effects that are statistically significant, except for Ella. Latifa and Ella seemed to have the most difficulty in applying their knowledge to unfamiliar words, presumably due to a lower degree of automation, especially considering that they both have a LD. Children with an enormous lack in reading competency often rely on the non-lexical route of the DRT, thus, trying to read while synthesizing (De Jong et al., 2012). This may be the reason for the rather slow-successive increase in the data of Latifa, Ella, and Lou (all with enormous deficits in pseudo-word reading) and, above all, the reason for any variability in the data. Overall, however, it can be said that the intervention had a transfer effect, but probably not enough, especially for weaker students. Nevertheless, the results are consistent with those of Mayer (2008), who found a greater effect on training words than on transfer words.

According to Kern and Manz (2004), goals (e.g., of an intervention) are socially invalid if they do not serve clients (e.g., students). In the current context, the survey of social validity showed that the support and its goals not only added value for the students, but that they also had a lot of fun. Thus, the intervention is a tool that can be used to increase reading skills beyond training words and that is also fully accepted by the students and is associated with fun and enjoyment. Motivation plays an enormous role, especially for students with low reading skills, and the racetracks as used here are both effective and enjoyable.

Limitations

Despite the encouraging results, some limitations of the study must be mentioned. First and foremost, it is difficult to generalize the findings, since we are only

dealing with a few children and whose characteristics are very similar. Second, restrictions due to the corona pandemic prevented students from different classes to be mixed. Thus, the children from third and fourth grade were separated. It would have been preferable to have some third and fourth graders in each group. However, fortunately, there was only a one-year grade difference; besides, there was a difference in reading ability between the third- and fourth-grade children. For example, some of the children scored higher numbers of correctly read words already in Phase A – especially Jim and Latifa – making it difficult to properly assess the effectiveness of the intervention because Phase B is capped with the number of words. This was particularly evident with Jim regarding the training words. However, in each case, a pool of 70 words was involved and 20 words were always randomly drawn – and not always the same 20 words. Jim achieved the maximum number almost every time in Phase B. Another limitation stems from the fact that the intervention consists of several method parts, and it is impossible to determine which part worked and how. However, since the intervention is very easy to use, it can be implemented well as a method package in schools if you know that the package works well. Furthermore, the children came from different language backgrounds and, as a result, their mother tongue may also have exerted an influence. Yet, Lou, the only native speaker of German, did not show any clear difference in the results compared to her classmates with German as a second language. Moreover, only the first group included two students with LD. Here it would also make sense that at least one child with LD would be represented in each group. Nevertheless, there are no clear differences between the students with and without LD.

Recommendations for Future Research

In future studies, the method package should be evaluated on a larger sample to yield stronger validity. In addition, the individual method aspects could be evaluated separately in a group or single-case design (e.g., A-B-BC plan) to determine the effects of the different aspects. It would also be interesting to find out whether participants' linguistic background makes a difference – that is, whether there is a clear difference between native speakers of German and students with German First Language (L1). Since some children in the study had severe problems in reading, and thus had difficulty memorizing the words as whole words (De Jong et al., 2012), it would make sense to reduce the number of training clusters and words for those children while at the same time increasing the frequency of occurrence of the clusters and words.

Conclusion

Overall, this single-case study showed that reading racetracks may be combined in such a way that it has not only an effect on trained words but possibly also on untrained words. For this purpose, it is sufficient to use direct instruction in frequently occurring letter clusters before implementing the racetracks and to practice them automatically. Thus, the racetracks, which have already shown very positive effects on reading in many studies (e.g., Barwasser et al., 2021a; Barwasser et al., 2021b; Grünke, 2019; Sperling et al., 2019) may be described as an even more powerful tool if combined appropriately for struggling readers with behavioral problems, with and without learning disabilities.

References

Alzahrani, T., & Leko, M. (2018). The effects of peer tutoring on the reading comprehension performance of secondary students with disabilities: A systematic review. *Reading & Writing Quarterly*, *34*(1), 1–17. https://doi.org/10.1080/10573569.2017.1302372

Barwasser, A., Urton, K., & Grünke, M. (2021a). Effects of a peer-tutorial reading racetrack on word fluency of secondary students with learning disabilities and emotional behavioral disorders. *Frontiers in Psychology, 12* (671385). https://doi.org/10.3389/fpsyg.2021.671385

Barwasser, A., Urton, K., Grünke, M., Sperling, M., & Coker, D. Jr. (2021b). 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

Bertin, J. C., & Narcy-Combes, J. (2007). Monitoring the learner – who, why and what for? *Computer Assisted Language Learning, 20*(5), 443–457. https://doi.org/10.1080/09588220701746021

Bos, W., Valtin, R., Hußmann, A., Wendt, H., & Goy, M. (2017). IGLU 2016: Wichtige Ergebnisse im Überblick [Important results at a glance]. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, & E.-M. Lankes et al. (Eds.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich [Reading competences of primary school children in Germany in international comparison]* (pp. 13–28). Waxmann.

Burns, M. K. (2007). Reading at the instructional level with children identified as learning disabled: Potential implications for response-to-intervention. School Psychology Quarterly, *22*(3), 297-313.

Coltheart, M. (2005). Modeling reading. The dual-route approach. In M. J. Snowling, & C. Hulme (Eds.), *The science of reading: A handbook* (pp.6–23). Blackwell. https://doi.org/10.1002/9780470757642.ch1

Crowley, K., McLaughlin, T., & Kahn, R. (2013). Using direct instruction flashcards and reading racetracks to improve sight word recognition of two elementary students with autism. *Journal of Developmental and Physical Disabilities*, *25*(3), 297–311. https://doi.org/10.1007/s10882-012-9307-z

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. https://doi.org/10.1080/09297049.2011.625354

Dyslexia Action. (2017). *Avoid dyslexia discrimination in the workplace.* https://dyslexiaaction.org.uk

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

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, *18*(1), 5–21. https://doi.org/10.1080/10888438.2013.819356

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, *53*(3), 199–208. https://doi.org/10.1177/001440298605300301

Geary, D. C., Hoard, M. K., Nugent, L., Ünal, Z. E., & Scofield, J. E. (2020). Comorbid learning difficulties in reading and mathematics: The role of intelligence and in-class attentive behavior. *Frontiers in Psychology*, *11*, 572099. https://doi.org/10.3389/fpsyg.2020.572099

Grabe, W. (2010). Fluency in reading – Thirty-five years later. *Reading in a Foreign Language*, *22*(1), 71–83. http://nflrc.hawaii.edu/rfl

Grigoryan, A. (2020). Major stages of reading skills development. *Armenian Journal of Special Education*, *1*(1), 42–51. https://doi.org/10.24234/se.2020.1.1.71

Grünke, M. (2019). The effects of reading racetracks on the sight word recognition of four elementary school students with learning difficulties. *International Electronic Journal of Elementary Education, 11*(4), 291–297. https://doi.org/10.26822/iejee.2019450787

Grünke, M., & Barwasser, A**.** (2019). Enhancing sight word fluency of second-language elementary students through reading racetracks. *International Journal of Technology and Inclusive Education*, *8*, 1373–1378.

Gunter, P. L., Miller, K. A., Venn, M. L., Thomas, K., & House, S. (2002). Self-graphing to success: Computerized data management. *Teaching Exceptional Children*, *35*(2), 30–34. https://doi.org/10.1177/004005990203500204

Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology*, *92*(2), 331–341. https://doi.org/10.1037/0022-0663.92.2.331

Harn, B. A., Stoolmiller, M., & Chard, D. J. (2008). Measuring the dimensions of alphabetic principle on the reading development of first graders: The role of automaticity and unitization. *Journal of Learning Disabilities*, *41*(2), 143–157. https://doi.org/10.1177/0022219407313585

Hasselhorn, M., & Gold, A. (2017). *Paedagogische Psychologie - Erfolgreiches Lernen und Lehren [Educational psychology - Successful learning and teaching]* (4th ed.). Kohlhammer.

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

Kern, L., & Manz, P. (2004). A look at current validity issues of school-wide behavior support. *Behavioral Disorders*, *30*(1), 47–59. https://doi.org/10.1177/019874290403000102

Knoepke, J., Richter, T., Isberner, M., Naumann, J., & Neeb, Y. (2014). Phonological recoding, orthographic decoding, and comprehension skills during reading acquisition. *Zeitschrift für Erziehungswissenschaft*, *17*(3), 447–471. https://doi.org/10.1007/s11618-014-0560-z

Kostewicz, D. E., Kubina, R. M., Selfridge, K. A., & Gallagher, D. L. (2016). A review of fixed fluency criteria in repeated reading studies. *Reading Improvement*, *53*(1), 23–33.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education*, *34*(1), 26–38. http://doi.org/10.1177/0741932512452794

Lee, J., & Yoon, S. (2017). The effects of repeated reading on reading fluency for students with reading disabilities. *Journal of Learning Disabilities*, *50*(2), 213–224. https://doi.org/10.1177/0022219415605194

Lerner, J., & Johns, B. (2011). *Learning disabilities and related mild disabilities: Characteristics, teaching strategies, and new directions* (11th ed.). Houghton Mifflin.

Macdonald, S. J., Deacon, L., & Merchant, J. (2016). "Too far gone": Dyslexia, homelessness, and pathways to drug use and dependency. *Insights into Learning Disabilities*, *13*(2), 117–134.

Marinak, B. A., & Gambrell, L. B. (2008). Intrinsic motivation and rewards: What sustains young children's engagement with text? *Literacy Research and Instruction*, *47*(1), 9–26. https://doi.org/10.1080/19388070701749546

Marinus, E., de Jong, P., & van der Leij, A. (2012). Increasing word-reading speed in poor readers: No additional benefits of explicit letter-cluster training. *Scientific Studies of Reading*, *16*(2), 166–185. https://doi.org/10.1080/10888438.2011.554471

Martin-Chang, S. L., & Levy, B. A. (2005). Fluency transfer: Differential gains in reading speed and accuracy following isolated word and context training. *Reading and Writing*, *18*, 343–376. https://doi.org/10.1007/s11145-005-0668-x

Mayer, A. (2008). *Phonologische Bewusstheit, Benennungsgeschwindigkeit und automatisierte Leseprozesse [Phonological awareness, naming speed and automated reading processes]*. Shaker.

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

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.

Morris, D., & Perney, J. (2018). Using a sight word measure to predict reading fluency problems in grades 1 to 3. *Reading & Writing Quarterly*, *34*(4), 1–11. https://doi.org/10.1080/10573569.2018.1446857

Mueller, B., Richter, T., Križan, A., Hecht, T., & Ennemoser, M. (2015). Word recognition skills moderate the effectiveness of reading strategy training in Grade 2. *Learning and Individual Differences*, *40*(5), 55–62. https://doi.org/10.1016/j.lindif.2015.05.001

Musti-Rao, S., Lo, Y., & Plati, E. (2015). Using an iPad® app to improve sight word reading fluency for at-risk first graders. *Remedial and Special Education*, *36*(3), 154–166. https://doi.org/10.1177/0741932514541485

Oakes, W. P., Mathur, S. R., & Lane, K. L. (2010). Reading interventions for students with challenging behavior: A focus on fluency. *Behavioral Disorders*, *35*(2), 120–139. https://doi.org/10.1177/019874291003500204

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data (PAND): An alternative to PND. *The Journal of Special Education*, *40*(4), 194–204. https://doi.org/10.1177/00224669070400040101

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification*, *35*(4), 303–322. https://doi.org/10.1177/0145445511399147

Rasinski, T. V., Rupley, W. H., Paige, D. D., & Nichols, W. D. (2016). Alternative text types to improve reading fluency for competent to struggling readers. *International Journal of Instruction*, *9*(1), 163–178. https://doi.org/10.12973/iji.2016.9113a

Sáenz, L. M., McMaster, K. L., Fuchs, D., & Fuchs, L. S. (2007). Peer-assisted learning strategies in reading for students with different learning needs. *Journal of Cognitive Education and Psychology*, *6*(3), 395–410. https://doi.org/10.1891/194589507787382089

Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007). *Reading interventions for adolescent struggling readers: A meta-analysis with implications for practice*. RMC Research Corporation, Center on Instruction.

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

Sideridis, G. D. (2002). Goal importance and students at risk of having language difficulties: An under explored aspect of student motivation. *Journal of Learning Disabilities*, *35*(4), 343–356. https://doi.org/10.1177/00222194020350040501

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. https://doi.org/10.1177/0022219411402691

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.

Sutherland, K., & Snyder, A. (2007). Effects or reciprocal peer tutoring and self-graphing on reading fluency and classroom behavior of middle school students with emotional or behavioral disorders. *Journal of Emotional and Behavioral Disorders*, *15*(2), 103–118. https://doi.org/10.1177/10634266070150020101

Tressoldi, P., Vio, C., & Iozzino, R. (2007). Efficacy of an intervention to improve fluency in children with developmental dyslexia in a regular orthography. *Journal of Learning Disabilities*, *40*(3), 203–209. https://doi.org/10.1177/00222194070400030201

van Keer, H., & Vanderlinde, R. (2010). The impact of cross-age peer tutoring on third and sixth graders' reading strategy awareness, reading strategy use, and reading comprehension. *Middle Grades Research Journal*, *5*(1), 33–45.

Vaughn, S., Levy, S., Coleman, M., & Bos, C. S. (2002). Reading instruction for students with LD and EBD: A synthesis of observation studies. *The Journal of Special Education*, *36*(1), 2-13. https://doi.org/10.1177/00224669020360010101

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

Young, C., Rasinski, T., Paige, D., & Rupley, W. (2020). Defining fluency. Finding the missing pieces for reading fluency. *Literacy Today*, *37*(6), 32–33.

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.

Zavala, E., & Cuevas, J. (2019). Effects of repeated reading and rhyming poetry on reading fluency. *International Journal of Social Sciences & Educational Studies*, *6*(2), 64–87. https://doi.org/10.23918/ijsses.v6i2p64