Between Languages, Texts and Narratives:

Adaptive Literacy Support for Diverse Learning Needs in Second Language

Learners in Different Age Ranges

Inaugural dissertation

for the attainment of the doctorate degree

of the Human Sciences Faculty University of Cologne in accordance with the

doctorate regulations of 18.12.2018

submitted by

Janine Ilona Bracht

from Dormagen, Germany

in September 2025

Supervised by Ph.D., Professor Matthias Grünke

**UNIVERSITY
OF COLOGNE**

Assessment

Primary Assessor: Matthias Grünke, Ph.D., Professor, University of Cologne, Germany

Second Assessor: Stephen Macdonald, Ph.D., Professor, Durham University, United Kingdom

This dissertation was accepted by the Faculty of Human Sciences at the University of Cologne in December 2025 (doctoral committee decision dated October 20, 2010).

The dissertation was defended on December 17, 2025.

Acknowledgements

> "I can no other answer make, but thanks,
>
> And thanks, and ever thanks."
>
> (William Shakespeare, *Twelfth Night* 3.3.14f)

I extend my heartfelt thanks to my doctoral supervisor, Professor Matthias Grünke, who believed in me when I was just starting my academic journey. He created opportunities I could hardly have dreamed of – not least this Ph.D. – and opened the door for me into the world of special education (and, quite unexpectedly, baseball). He provided me with new perspectives and inspiring ideas that have already had a lasting impact on my academic career, even before the completion of this doctorate. Thank you for making this happen!

I would also like to express my sincere gratitude to Professor Stephen Macdonald for reviewing my work. It is both an honor and a joy that the very first international colleague I met at my first conference is now supporting me by kindly taking on the role of the second assessor for my dissertation.

My thanks also go to my team, who accompanied me on my path to completing this doctorate through countless inspiring discussions and an ever-motivating atmosphere. A special thank you goes to Dr. Susanne Hoff, who, particularly in the final stages, brought clarity to moments of confusion, lent a listening ear in times of doubt, and offered strength during phases of uncertainty.

I would like to dedicate my deepest gratitude to Dr. Anne Barwasser, who has supported me with great selflessness and friendship since the day we met. I have learned so much from her, and I look forward to the exciting projects we are planning together for the future.

I am also profoundly grateful to my family and friends for their understanding, trust, and support throughout my academic career. In particular, I thank my parents, whose unwavering support and unconditional belief in me have always given me the strength to pursue my goals with determination.

The journey to this dissertation has been exciting, inspiring, and full of surprises. It has been a learning experience that has left a lasting mark on me, both professionally and personally. It is a wonderful feeling to complete this chapter and to be able to embrace new challenges.

Abstract

"What do you read, my lord?
Words, words, words."
(William Shakespeare, *Hamlet* 2.2.191f)

While Hamlet may regard words as nothing more than words, for many students learning German as a second language they represent thresholds, and their combination into coherent texts poses a major hurdle. It is precisely at this point that the present dissertation is anchored. It investigates adaptive, evidence-based interventions that support learners at different developmental stages in progressing from word to meaning, from sentence to text, and ultimately to narration.

Against the backdrop of increasing heterogeneity and a steadily growing number of students with German as a second language, five studies are presented, conducted in secondary school, primary school, and preschool. Two of the studies evaluate an adapted German version of the Peer-Assisted Learning Strategies program with secondary students experiencing or at risk of reading difficulties and learning German as a second language. The results reveal significant gains in reading fluency and comprehension, as well as high acceptance of the intervention among both students and teachers. Another study with third- and fourth-grade students combined peer tutoring, story maps, and Self-Regulated Strategy Development. This approach led to significant improvements in reading and showed positive trends in writing.

At school entry, a narrative intervention proved effective in strengthening vocabulary, letter-sound fluency, and basic reading skills. In preschool, however, results showed that storytelling alone was insufficient; only when combined with direct instruction did significant progress emerge in grapheme-phoneme correspondence and vocabulary. Methodologically, the studies employed multiple-baseline single-case designs as well as an experimental group design. Social validity and feasibility were also systematically assessed.

In conclusion, the results are synthesized and implications for research and practice are outlined. Overall, the dissertation demonstrates that adaptive interventions—through peer tutoring, strategy training, motivational components, incidental and intentional learning—can effectively reduce barriers faced by students with special learning needs or German as a second language. In doing so, it provides practice-oriented guidance for fostering literacy within the German educational context while also contributing to the prevention of educational disadvantage.

Table of Contents

List of Abbreviations

ANCOVA	Analysis of Covariance
BAKO 1–4	Basic Skills for Reading and Spelling Test
BC-SMD	Between-Case Standardized Mean Difference
DCT	Dual Coding Theory
DESI	German English Student Performance International Study
DI	Direct Instruction
DRT	Dual Route Theory
ELFE-II	Reading Comprehension Test for First to Seventh Graders – Version II
GL2	German as a Second Language
GPC	Grapheme-Phoneme Correspondence
GraWo	Graz Vocabulary Test
GRR	Gradual Release of Responsibility
IQB	Institute for Quality Development in Education
ITRF-G	Integrated Teacher Report Form German
L1	First Language
L2	Second Language
LD	Learning Disabilities
LDL	Reading Progress Diagnostics
LSF	Letter-Sound Fluency
NAP	Non-Overlap of All Pairs
OECD	Organisation for Economic Cooperation and Development
PA	Phonological Awareness
PALS	Peer-Assisted Learning Strategies
PAND	Percentage of All Non-Overlapping Data
PEM	Percentage of Data Points Exceeding the Median
PEM-T	Percentage of Data Exceeding a Median Trend
PISA	Programme for International Student Assessment
PT	Peer-Tutoring
RCT	Randomized Controlled Trial
SCR	Single-Case-Research
SET 5–10	Language Assessment Test for Children Aged Between 5 and 10
SLRT-II	Further Development of the Salzburg Reading and Spelling Test
SRSD	Self-Regulated Strategy Development
SVR	Sachverständigenrat deutscher Stiftungen für Integration und Migration
SWRD	Specific Word Reading Difficulties
TEPHOBE	Test for the Acquisition of Phonological Awareness and Naming Speed
VSL	Progress Diagnostics of Reading Comprehension
WM	Working Memory
WS/ZF-R	Vocabulary Test and Number Sequence Test Revision
ZPD	Zone of Proximal Development

1. Introduction

Literacy encompasses a wide range of concepts that are closely tied to language and extends far beyond basic reading and writing. It includes additional linguistic components such as vocabulary, grammar, and precursor skills like phonological awareness (PA). These foundational abilities are acquired during early childhood and continue to develop throughout the school years (Castles et al., 2018). Literacy shapes our everyday lives, our educational trajectories, and eventually our professional lives—it is omnipresent and holds a central role in education. The more intensive and varied the contact with language, the richer the experiences accumulated, and the better the development of literacy competence in its broadest sense (Dong & Chow, 2022).

The Lexical Quality Hypothesis proposed by Perfetti and Hart (2002) suggests that high lexical quality leads to more efficient word recognition, better text comprehension, and, ultimately, more successful text production. For second language (L2) learners, however, this presents a particular challenge: without the necessary linguistic tools in the L2, the development of later reading and writing competence is substantially hindered (Kwakkel et al., 2024).

In light of global change and increasing mobility, the number of individuals acquiring a L2 continues to rise (Duarte et al., 2020). In Germany as well, the number of children and adolescents learning German as a second language (GL2) is steadily growing. According to the 2024 Microcensus (Statistisches Bundesamt, 2024a), approximately 42 percent of all students in Germany had a migration background. This presents a central challenge: understanding the specific needs of students learning GL2. This group is as diverse and heterogeneous as any other learning population. A migration background alone provides little information about individual linguistic biographies. What is decisive is whether and at what age a child was first exposed to the L2, the nature and quality of that exposure, and the additional support opportunities available. All of these factors substantially shape children's lexical quality in the L2 (Cadierno, 2022; Thordardottir & Plez, 2024).

At the same time, studies consistently demonstrate that students with a migration background often rank among the lower-performing groups in achievement assessments. For example, the 2022 The Programme for International Student Assessment (PISA) results showed that non-immigrant students scored, on average, 67 points higher in reading than their peers with a migration background (Organisation for Economic Cooperation and Development [OECD], 2023). With the increasing number of students learning GL2, the challenge of providing appropriate linguistic support becomes even more pressing. Moreover, students with a migration background are disproportionately affected by socioeconomic disadvantage: 42 percent

compared to 25 percent in the overall population (Anger et al., 2024; Grünke & Bracht, 2025; OECD, 2023). Such disadvantages are closely linked to linguistic difficulties, which stem from insufficient prerequisites for the comprehensive development of literacy skills.

Language opens doors—it is the foundation of social participation. Taking responsibility for language development therefore means helping those who, within the educational system, would otherwise be more likely to find doors closed to them. The aim of this work is to open those doors through language support, to provide assistance where it is needed, as individually as possible and as early as necessary.

In addition to the challenges, the advantages of multilingualism should not be over-looked. Multilingualism brings cognitive strengths, fosters a refined sensitivity to language, and can be strategically leveraged to facilitate the acquisition of additional languages (Yurtsever et al., 2023). Existing research shows that literacy support must always take into account the specific conditions of L2 acquisition (Kittle et al., 2024). The present work aims to contribute to this field by developing accessible interventions that address the needs of children at different ages and with diverse learning needs and first languages (L1)—recognizing their full diversity as learners of GL2.

From this context emerges the central research question of this dissertation: *How can adaptive, evidence-based literacy interventions be designed to meet the diverse linguistic and learning needs of L2 learners across different age groups—from early childhood to adolescence—in order to effectively strengthen their reading and writing skills as well as their foundational precursor abilities and vocabulary?*

This dissertation first addresses advanced literacy competence, namely reading and writing proficiency, as well as the barriers that learners face in these domains and the instructional approaches that research has identified as particularly promising. The first two empirical studies focus on a German-translated and adapted version of the Peer-Assisted Learning Strategies program (PALS) developed by Fuchs et al. (1997). The aim was to investigate the extent to which this intervention is effective for secondary school students, particularly those at risk for reading difficulties and L2 learners. The studies first examine students in the seventh grade and subsequently younger students in the fifth and sixth grades. The third study also centers on peer-tutoring (PT) in reading, but with younger learners in the third and fourth grade. The intervention employed PT methods in combination with story maps in order to improve both reading and writing competence.

Building on this, the dissertation then turns to early literacy competence. The focus lies on precursor skills and vocabulary—foundational components that are crucial even before the

development of advanced literacy (Castles et al., 2018). The final two studies therefore examine younger children, specifically first graders and even preschoolers, in order to assess the effectiveness of early interventions in language development. The goal is to provide timely support to students who struggle or who are learning GL2, thereby addressing potential barriers to literacy acquisition at the earliest possible stage.

Finally, conclusions are drawn from all studies, and implications for both research and practice are discussed. If targeted interventions succeed in meeting the needs of learners across different age groups and with diverse learning requirements—particularly in cases of learning difficulties and L2 acquisition—this opens up valuable pathways to address obstacles in language development as early as possible and with timely, appropriate support measures.

2. Advanced Literacy Support for Struggling Readers and German as a Second Language Learners

2.1 Text Comprehension and Production

Reading and writing are foundational components of everyday life and, above all, of formal schooling (Castles et al., 2018). These skills are introduced and practiced in elementary school and are then deepened and refined throughout a student's academic career. Typically, reading develops earlier than writing because receptive abilities are generally easier to acquire and place lower cognitive demands than expressive skills (Dixon et al., 2023; Nation, 2013). Accordingly, reading is often perceived as simpler and less demanding than independent writing.

Reading ability can be divided into two core domains: reading comprehension and reading fluency. Both are necessary for proficient reading, yet they stand in a distinct relation to each other (Kim et al., 2021b). It is possible to read a text fluently without grasping its meaning, just as it is possible to read slowly and laboriously while nonetheless constructing a solid understanding. Only the interplay of both components produces a competent reader. Reading competence may therefore be described as the ability to read words, sentences, or texts with fluency while comprehending their content (Castles et al., 2018; Kim et al., 2021b).

These two dimensions are essential not only at school but also in everyday contexts, whether understanding task instructions in other subjects or making sense of a user manual. The

development of reading begins in elementary school with vocabulary work and precursor skills, moves through word- and sentence-level reading, and culminates—typically by the end of elementary school—in text-level reading. This progression also prepares students for the transition to secondary school, which in the German system starts after grade four. From that point onward, core curricula already presuppose established reading competencies (Rohm et al., 2021). Delays in the development of reading leave little room for remediation and must be addressed swiftly, lest deficits compound and create substantial learning barriers across all subjects, including non-language domains. These gaps become especially apparent when students solve word problems in mathematics (Fuchs et al., 2019; Koponen et al., 2018). Strong reading skills open the door to disciplinary learning, whereas weak readers once again encounter a "closed door". To be able to read is, in effect, to be able to acquire knowledge.

Following Massler et al. (2022), reading fluency comprises four interrelated dimensions—coding accuracy, meaning the correct recognition of words; automatization, meaning recognition without letter-by-letter decoding; reading rate, meaning the efficient processing of words and multiword units; and prosody, meaning meaningful oral reading or appropriately stressed silent reading. Fluency is a necessary but not sufficient condition for comprehension (Kim et al., 2021b). Fluent reading reduces working-memory load and enables readers to allocate more cognitive resources to meaning (Kim et al., 2021b). For this reason, it is sensible to cultivate fluency and comprehension in tandem. Empirical studies report a moderate to substantial correlation between the two (Kim et al., 2021b; Sappok et al., 2020). International comparative studies further indicate that, by the end of elementary school, meaning-oriented reading comes to the fore (Kang & Shin, 2019; Rohm et al., 2021); fluency facilitates comprehension but does not guarantee it. As students progress through the grades, texts in educational settings become more complex, which argues for sustained support into the upper grades. Reading difficulties should therefore be identified as early as possible and addressed with targeted intervention.

The Simple View of Reading (Gough & Tunmer, 1986) clarifies this logic by modeling reading comprehension as the product of decoding and linguistic comprehension. If either component is absent, the product is zero and comprehension does not occur. Students who can accurately recognize words but lack the linguistic knowledge to interpret them understand no better than students with strong oral language who nevertheless fail at decoding, as in dyslexia. Numerous studies have corroborated this model and offer a pragmatic basis for instruction: when difficulties lie in decoding, precursor skills such as PA should be explicitly trained (Castles et al., 2018; Rehfeld et al., 2022); when the bottleneck is linguistic comprehension,

instruction should focus on vocabulary, comprehension strategies, and background knowledge (Wright & Cervetti, 2017)—especially salient during the transition to secondary school. When both components are weak, a combined approach is required.

It is indisputable that students who read fluently and possess robust linguistic comprehension read better. But does better reading also make for better writing? Research indicates a reciprocal relationship: reading and writing mutually support one another and draw on shared linguistic foundations and cognitive processes (Graham, 2020; Kim & Zagata, 2024; Schoonen, 2019). Extensive reading not only strengthens reading competence but also refines linguistic sensitivity, expands vocabulary, and deepens understanding of linguistic structure—all central prerequisites for composing one's own texts. Numerous meta-analyses, such as Graham and Hebert (2011), confirm this association.

At the same time, good readers are not automatically good writers. Writing entails additional linguistic and cognitive processes: not merely recognizing structures but deploying them independently and accurately (Ahmed et al., 2021; Galbraith & Baaijen, 2018). Reading and writing are thus distinct competencies that both require targeted support, even as they benefit from common underpinnings. Particularly effective are approaches that integrate the two—for example, having students process texts through writing by summarizing, revising, and reflecting—practices that have been shown to improve text comprehension as well (Graham, 2020; Kim & Zagata, 2024; Stevens et al., 2018).

With the transition to secondary school, the emphasis in literacy development shifts progressively from the word and sentence levels to the text level. Whereas elementary instruction often foregrounds discrete linguistic elements, secondary students must extract information from more complex genres, recognize and apply textual structures, and meet demands that place greater burdens on cognitive resources (Kraal et al., 2018). Research describes these as higher-level processes: identifying a text's central ideas, drawing inferences, interpreting and evaluating content, and developing an informed stance based on what has been read (Allen & McNamara, 2020; Elleman, 2017).

This trajectory underscores the need to continue literacy instruction systematically beyond elementary school and into secondary education. Support cannot stop with basic decoding; it must take the holistic text as its focus. The challenges, approaches, and specific instructional strategies pertinent to this endeavor are discussed in detail in the chapters that follow.

2.2 Literacy Difficulties and German as a Second Language

The broad spectrum of literacy skills inevitably leads to the possibility that students may experience difficulties in individual areas of language development or even across all domains. In the context of supporting L2 learners, a central diagnostic dilemma must be considered: whether challenges in literacy development stem from specific linguistic deficits such as dyslexia or from learning-related impairments such as learning disabilities (LD), or whether they merely reflect delays in L2 acquisition. This dilemma carries the risk of overidentifying children and adolescents who do not, in fact, have severe learning difficulties, or underidentifying those who genuinely require more comprehensive support (Hall, 2009; Taha et al., 2022; Zhang & Wang, 2023).

Research shows that students learning GL2 are more frequently misdiagnosed with learning disorders (Brandenburg et al., 2016). One major reason is the lack of adequate testing options in the students' home languages, as well as the absence of norms for GL2 learners in standardized assessments. For example, 25–30% of third graders with GL2 received a diagnosis of reading and spelling difficulties, almost twice as many as students with German as their L1 (14–18%) (Brandenburg et al., 2016). When adjusted norms are applied, however, the diagnostic rates of L1 and L2 learners converge significantly. This demonstrates the clear risk of over-diagnosis among GL2 learners and highlights the need to distinguish genuine LD from language-related barriers. A promising approach would be the use of comprehensive multilingual assessments or nonverbal testing procedures (Brandenburg et al., 2016).

If a child already shows literacy-related difficulties in their L1, Cummins' interdependence hypothesis (1979a) becomes relevant. This hypothesis emphasizes the link between L1 and L2, noting that language skills are grounded in common cognitive foundations. Strategies and metacognitive abilities developed in L1 can be transferred positively to L2. Conversely, deficits in L1 may significantly hinder L2 acquisition. Vocabulary is especially critical, since a limited lexicon almost always represents the greatest obstacle for L2 learners (Zhang & Zhang, 2022).

Another central factor is the distinction between everyday language (Basic Interpersonal Communication Skills) and academic language (Cognitive Academic Language Proficiency), as proposed by Cummins (1979b). Everyday language is typically acquired quickly, since it is practiced in direct interactions with peers, is highly context-dependent, and does not require linguistic precision. Academic language, by contrast, is tied to school contexts, more formalized, and demands greater accuracy. It is more abstract and less anchored in daily life, which makes its acquisition considerably slower (Phillips Galloway et al., 2020). While students can

usually acquire everyday language within about two years, mastering academic language often takes up to seven years (Lange & Gogolin, 2010). Because school systems rarely account for these timelines, significant learning gaps arise that are difficult to close. Even when everyday language is mastered, insufficient development of academic language frequently leads to major limitations in academic achievement. Several studies demonstrated that academic language is directly correlated with school success (Cruz Neri et al., 2021; Phillips Galloway & Uccelli, 2019).

In addition to linguistic barriers, cognitive resources—particularly working memory (WM)—play a central role. WM enables the simultaneous processing of multiple pieces of information and is heavily taxed during both reading and writing (De Vita et al., 2021; Nouwens et al., 2021). When its capacity is already consumed by basic processes such as decoding, word recognition, and sentence construction, little remains for higher-order demands such as coherence, precision, creativity, or structured planning. According to models of reading and writing research (Hayes & Flower, 1980; Kellogg, 1996; Nouwens et al., 2021), however, these simultaneous processes are characteristic of skilled readers and writers: they plan ahead, organize content, and focus more strongly on meaning and linguistic appropriateness.

For less experienced L2 learners, such processes are often not yet automated, which quickly results in WM overload. Because linguistic routines are lacking, additional cognitive resources must be allocated to decoding and comprehension (Li, 2023). This leaves little capacity for depth of content or coherence in reading and writing. The result is shorter and less structured texts, slow or halting reading, or repeated rereading of individual passages (Grabe & Kaplan, 1996). Slowed processing further impairs comprehension, since the focus on basic tasks narrows the ability to grasp the overall text (Junk-Deppenmeier & Jeuk, 2015).

Slowed writing also often reduces quality, and in time-pressured school contexts it adds to frustration and fear of making mistakes (Grünke & Bracht, 2025; Rasool et al., 2023; Skar et al., 2022). These negative experiences again burden WM, lower motivation, and reinforce feelings of reduced self-efficacy. This creates a vicious cycle: reduced motivation leads to less practice, less practice intensifies cognitive overload, and this in turn generates further frustration.

It is important to emphasize that such difficulties are not related to intelligence or lack of willingness to learn. Instead, they are closely tied to psychosocial factors such as self-efficacy, motivation, and resilience, which are both shaped by literacy-related challenges and themselves influence reading and writing processes (Sehlström et al., 2023; Toste et al., 2020). Once

again, a circular process becomes apparent, one that needs to be guided through targeted and sensitive educational support.

According to Bandura (1978), self-efficacy refers to an individual's belief in their ability to successfully complete a task through their own efforts. This belief is strongly shaped by prior experiences, with recurring patterns playing a particularly significant role in shaping a student's self-perception. Repeated failure in reading and writing or the constant experience of poor grades inevitably undermines confidence in one's abilities. Conversely, positive experiences strengthen trust in one's own competence (Talsma et al., 2018).

For L2 learners, the risk of low self-efficacy is particularly high, since—as outlined above—they disproportionately struggle with linguistic skills. They operate within an educational system conducted entirely in the target language and are therefore continually confronted with their linguistic limitations. This fosters a strong deficit focus. Constant comparison with L1 learners further diminishes their perception of their own abilities (Grosjean, 1985).

Research shows that students with low self-efficacy invest less effort and reach their motivational limits more quickly, as they do not believe their efforts will yield positive outcomes (Schunk & DiBenedetto, 2020). Past failures often lead to avoidance behaviors, reduced practice, and consequently to entrenched difficulties (Tahmouresi & Papi, 2021). Low self-efficacy thus results in lower motivation and weaker performance—those who read little remain weak readers and become increasingly demotivated.

For this reason, support measures must not only address linguistic barriers but also take into account the often already diminished sense of self-efficacy. The challenges L2 learners face in literacy are as diverse as the learners themselves. These challenges must be recognized in their full complexity and explicitly considered in instructional and support frameworks. Approaches to comprehensive, holistic, and ambitious support that address this complexity will be presented in the following chapter.

2.3 Enhancing Literacy Competence

The promotion of literacy competence requires theoretical foundations and methods that are effective for general learning. These include fostering independent learning through problem-solving processes, reducing WM load through supportive tools, overcoming barriers through collaborative assistance, and providing targeted adaptations for reading and writing, such as reading strategies or writing planning aids.

A central theoretical foundation is Vygotsky's sociocultural learning theory (1978), which posits that learning always occurs within a social and cultural context and is more successful when it is collaborative. Of particular importance is the Zone of Proximal Development (ZPD), which refers to the range of skills that learners cannot yet master independently but can achieve with appropriate support (van de Pol et al., 2019). Crucially, assistance must be adapted to the learner's developmental level: it should not be so extensive that the learner remains passive, nor so minimal that the task remains unmanageable. A useful analogy is learning to ride a bicycle: training wheels serve as necessary supports at the beginning but must eventually be removed to enable independent riding.

Learning tasks should therefore be carefully selected to fall within, rather than below or above, the ZPD. Alongside individual support, cooperative learning and interaction play a central role, as collaborative work facilitates learning within the ZPD (Tenenbaum et al., 2020). One example is the group contingency procedure, in which students share responsibility for a group task and pursue a common goal. Shared responsibility can encourage students to take their own learning more seriously and to motivate one another (Tenenbaum et al., 2020).

A similar sense of responsibility emerges in PT. Here, students work in pairs, typically pairing a stronger with a weaker learner, with the stronger student serving as tutor (Sáenz et al., 2005; Thurston et al., 2021). Both benefit from the exchange: the tutor consolidates knowledge while the tutee receives targeted support. PT can also be effective among students of similar ability, as the roles can be flexibly exchanged (Gazula et al., 2017).

Building on these principles, the PALS program (Fuchs et al., 1997) was developed. It provides a structured framework for paired reading, with defined roles and materials. The strategy steps include partner reading, error correction, paragraph summarization, and prediction relay. PALS has been widely tested in the Anglo-American context and has demonstrated strong effectiveness (Fuchs et al., 1997, 2020, 2021; Lee et al., 2023). However, it has rarely been applied with L2 learners or students with LD (Sáenz et al., 2005; Spörer et al., 2009), having been studied primarily in heterogeneous groups without specific support needs.

Another approach is reciprocal teaching (Palincsar & Brown, 1984), in which teachers model metacognitive strategies that are then practiced in small groups or pairs with rotating roles. Unlike PALS, reciprocal teaching emphasizes reflecting on texts through metacognitive strategies rather than following fixed reading steps.

Strategies are widely regarded as especially suitable for literacy development because of their long-term impact (Chen et al., 2021; Gillespie Rouse & Sandoval, 2018; Sanders, 2020; Sanders et al., 2019): learners acquire tools that help them manage difficulties independently.

A prominent example is Self-Regulated Strategy Development (SRSD), developed by Harris and Graham (1996). This approach combines strategy training with self-regulation and follows a six-step process: (1) develop background knowledge, (2) discuss, (3) model, (4) memorize, (5) guided practice, and (6) independent performance. Through self-regulatory processes such as goal-setting and self-motivation, learners can continue applying these strategies over the long term, even without close teacher supervision.

The SRSD approach is highly flexible and can be adapted to a variety of tasks, including planning narrative texts, preparing argumentative essays, or systematically managing reading processes. International meta-analyses report large effect sizes for SRSD interventions ranging from .59 to 1.04 (Graham & Harris, 2018; Kim et al., 2021a).

To make practice phases especially effective, scaffolding is essential. Scaffolding refers to instructional supports that guide learners toward the next developmental stage within their ZPD (van de Pol et al., 2019). This can take various forms: cognitive (e.g., structuring tasks in advance), linguistic (e.g., providing vocabulary pools or sentence starters), motivational (e.g., encouragement and positive feedback), or social (e.g., cooperative learning structures). A particularly effective form of cognitive scaffolding is the use of graphic organizers (Sencibaugh, 2007; Sun et al., 2021). These tools provide learners with support through writing frames, checklists, or visual diagrams that help them structure complex information. By externalizing planning and organizing processes before a full text is produced, graphic organizers reduce the cognitive demands placed on WM (Krieglstein et al., 2022).

Story maps provide a concrete example. They highlight key elements of a narrative, offering fields for title, characters, setting, problems, and solutions. Such structures encourage students to engage with texts actively rather than consume them passively. This not only strengthens reading comprehension but also promotes an increasing awareness of narrative structures (Bogaerds-Hazenberg et al., 2020). Empirical studies show that story maps are particularly beneficial for struggling readers (Cure et al., 2020). Overall, graphic organizers are considered evidence-based interventions and are explicitly recommended by the U.S. National Reading Panel (2000).

In summary, the range of possibilities for promoting literacy is diverse and must always be tailored to the needs of the specific learner group. Vygotsky's Sociocultural Learning Theory (1978) provides a robust foundation for designing cooperative and motivating learning processes. Strategy-based approaches such as SRSD equip learners with sustainable tools, while graphic organizers relieve WM and provide structure. Together, these approaches serve as

effective "training wheels" that help students overcome the next challenges in their literacy development.

2.4 Publications

The theoretical framework makes clear that comprehensive literacy support must address multiple linguistic dimensions while at the same time reducing cognitive load and fostering motivation. At the same time, there is a research gap in the German-speaking context, particularly with respect to L2 support for learners facing either learning difficulties or language barriers or even both. While cooperative and strategy-oriented approaches have already been shown to be effective in numerous international studies (Okkinga et al., 2018; Tenenbaum et al., 2020), there is still little research on how these interventions can be adapted to the specific needs of L2 learners in German-speaking schools, both in terms of linguistic particularities and within the structures of the German school system.

Students learning GL2 have rarely been at the center of empirical intervention research, even though they face distinctive linguistic challenges that must be clearly differentiated from general LD and that require targeted support. The transferability of existing programs such as the PALS program or scaffolding tools such as story maps into the German-speaking context, as well as the systematic examination of their suitability for L2 learners, therefore represents a central focus of the following studies. In doing so, they not only expand the existing body of research but also make an important contribution to evidence-based literacy support in the German educational system, at a time when such questions are more pressing than ever.

2.4.1 Summary Article 1 *(peer-reviewed)*

Hertel, S., Bracht, J., Calhoon, M. B., Grünke, M., & Barwasser, A. (2024). Effects of an adapted peer-assisted learning strategies reading programme on reading fluency and reading comprehension of secondary students with or at-risk for reading disabilities. *European Journal of Special Needs Education*, 1–17. https://doi.org/10.1080/08856257.2024.2402166

The effectiveness of PT approaches has been well established in educational research, as evidenced by numerous meta-analyses (Bowman-Perrott et al., 2013; Moeyaert et al., 2021). In the domain of reading, collaboration in pairs or small groups has consistently been shown to enhance reading competence. In particular, the PALS Reading 2–6 program developed by Fuchs

et al. (1997) has demonstrated promising results in improving text-level reading comprehension across multiple studies (Fuchs et al., 1997, 2020, 2021; Lee et al., 2023; Sáenz et al., 2005). Against this backdrop, the present study sought to examine whether an adapted version of PALS, culturally and linguistically tailored for the German context, could support secondary school students with specific word reading difficulties (SWRD). Until now, PALS has only been adapted for use in secondary education by Spörer et al. (2009) in Germany. However, the present study has shortened the program again in a different way and tested it in a single-case study design with students with SWRD or at risk of developing them.

Method

The study was conducted at a secondary school in Germany with seventh-grade students. Twelve adolescents initially participated in the intervention, but due to extensive absences related to COVID-19, only eight were included in the final analysis. All exhibited low reading comprehension performance (percentile rank < 25) on the Progress Diagnostics of Reading Comprehension (VSL; Walter, 2013) and Reading Comprehension Test for First to Seventh Graders – Version II (ELFE-II; Lenhard et al., 2020). Participants were 13 to 14 years old, represented diverse migration backgrounds, and some had additional special educational needs in the areas of learning or speech and language.

The intervention was implemented in groups of four students, with pair-based collaboration within each group, resulting in three groups overall. The study employed a multiple-baseline across participants design following an AB structure (Ledford & Gast, 2024): Phase A (baseline) without intervention, and Phase B (intervention with the adapted PALS program). Sessions were held three times per week, lasting 45 minutes each, with a total of five to seven intervention sessions depending on baseline length.

Screening instruments included the ELFE-II test (reading comprehension at word, sentence, and text levels; Lenhard et al., 2020), the VSL for reading accuracy (Walter, 2013), the Reading Progress Diagnostics (LDL; Walter, 2009) for reading fluency, and the Integrated Teacher Report Form German (ITRF-G) for behavioral and academic assessment (Volpe et al., 2018). Dependent variables were reading comprehension (VSL) and reading fluency (LDL), measured repeatedly across phases.

The adapted PALS program comprised four core reading strategies: partner reading (mutual error correction during oral reading), retelling (summarizing sections of text), paragraph shrinking (identifying main ideas), and the prediction rally (making predictions about text progression). Students received a materials package including books, role descriptions, rule cards, and a reinforcement system in which points were collected with marbles. Unlike the

original PALS program, the roles of "Coach" and "Athlete" alternated each session, eliminating the fixed hierarchy between first- and second-reader. After five introductory sessions, the program was fully implemented beginning with session six.

The intervention and data collection were carried out by trained graduate students in special education. Treatment fidelity was monitored through independent observation, yielding an interrater reliability of 91% (Sanetti et al., 2021). Social validity was assessed at the end of the program via a 16-item student questionnaire (acceptance and perception) and an 8-item teacher questionnaire (effects in regular class, motivation, enjoyment).

Results

Findings indicated that seven of eight students demonstrated significant gains in reading fluency. The Between-Case Standardized Mean Difference (BC-SMD; Chen et al., 2023) revealed a significant effect of 1.21 (CI95 [0.53, 1.89]), which remained stable during follow-up. Reading comprehension also improved among all students, including those with elevated behavioral scores on the ITRF (score > 13).

Social validity ratings were predominantly positive. Students valued peer error correction and the marble-based reward system, with girls evaluating the program more favorably than boys. Teachers also reported positive impressions, noting increased motivation and engagement. However, no direct correlation was found between subjective ratings and objectively measured learning gains.

Discussion

Overall, the study demonstrated that the adapted German version of PALS yielded significant improvements in reading fluency and also enhanced reading comprehension. The most substantial progress was observed among students with stronger baseline word-reading skills, as reflected in the ELFE-II screening. Importantly, even students with behavioral difficulties benefitted from the intervention, suggesting that improvements in reading can occur independently of behavioral challenges.

The program was positively evaluated, particularly by female students, with enjoyment and reinforcement mechanisms highlighted as strengths. Notably, subjective perceptions did not consistently align with objective performance gains, underscoring the complexity of motivational and affective dimensions in intervention research.

The study is strengthened by its robust single-case design with staggered baselines, enabling attribution of effects to the intervention (Ledford & Gast, 2024; Kratochwill et al., 2013). The inclusion of both objective performance data and subjective evaluations further enhances

the validity of findings. The adapted PALS program proved feasible and practical for implementation in everyday school contexts.

As with all single-case designs, generalizability is limited. COVID-19-related absences further reduced sample size, and the intervention duration was considerably shorter than in the original PALS program (Fuchs et al., 1997). Moreover, no data on long-term sustainability of effects are yet available.

Future research should replicate these findings with larger groups, such as entire classes of students with reading difficulties. Investigating teacher-led implementation of PALS would also provide valuable insights into ecological validity. The social validity scale requires further validation, and follow-up studies should examine the durability of effects over time.

In conclusion, the adapted German version of PALS appears to be an effective, peer-based reading intervention that enhances both fluency and comprehension. The program was well accepted by students and teachers and represents a practical option for classroom use in heterogeneous learning groups, including students from migrant backgrounds, for whom the intervention may help mitigate language-related disadvantages.

2.4.2 Summary Article 2 *(peer-reviewed)*

Bracht, J., Hoff, S., Grünke, M., & Barwasser, A. (2025_accepted). Enhancing reading competencies of German as a second language learners through an adapted peer-assisted learning strategies reading programme. *Insights into Learning Disabilities*.

The previous study has demonstrated that peer-tutored reading programs such as PALS-Reading (Fuchs et al., 1997) can be successfully implemented with older secondary students, with particular benefits for learners of GL2. This raised the question of whether the program is equally effective with younger students, particularly during the transition from elementary to secondary school. In the German school system, this transition usually occurs between grades 4 and 5 and represents a sensitive developmental phase. L2 learners, in particular, face distinct linguistic and academic challenges during this period (Dollmann & Weißmann, 2020). The aim of the present study was therefore to examine the effectiveness of PALS with younger students at this transitional stage and suitability for GL2 learners.

Method

The study employed a single-case multiple-baseline design with three phases: baseline (A), intervention (B), and follow-up (E) (Ledford & Gast, 2024). Data were collected at all phases approximately three times per week over an eight-week period.

The study was conducted in grades 5 and 6 at two secondary schools in North Rhine-Westphalia, Germany, including preparatory classes for L2 learners. From an initial sample of 43 students, eight were included in the final study sample. Seven were L2 learners, and one was a native German speaker who participated at the teacher's request. Inclusion criteria were poor reading comprehension (percentile rank < 25 on the ELFE-II test at the text level and VSL < 15) as well as vocabulary and language deficits identified with the Graz Vocabulary Test (GraWo; Seifert et al., 2017) and Language Assessment Test for Children Aged Between 5 and 10 (SET 5–10; Petermann, 2018).

To assess reading competence, a screening battery was administered consisting of ELFE-II (word-, sentence-, and text-level comprehension; Lenhard et al., 2020), the VSL for reading accuracy (later used as a dependent variable; Walter, 2013), the LDL for fluency (Walter, 2009), the GraWo for receptive vocabulary (Seifert et al., 2017), and the SET 5–10 for expressive vocabulary (Petermann, 2018). In addition, the ITRF-G teacher questionnaire on social behavior and academic performance was used (Volpe et al., 2018). Reading accuracy (VSL) and reading fluency (LDL) were defined as dependent variables.

The intervention again was an adapted German version of the U.S. PALS Reading program for grades 2–6 (Fuchs et al., 1997). The four core steps—partner reading, retelling, paragraph shrinking, and prediction rally—were implemented throughout the intervention. Students worked in dyads within small groups, alternating between the roles of "Coach" and "Athlete." Age-appropriate youth books in simplified language were used as reading materials. A point system with stickers and certificates served as an additional motivational component.

Several adaptations were made to the original PALS program (Fuchs et al., 1997). The duration was shortened from 15 to 6 weeks to facilitate implementation in everyday school practice; the number of introductory sessions varied flexibly between 6 and 12 depending on the time required to master each step. Materials were translated and culturally adapted. Introductory short-story exercises, included in the original version, were omitted, and students began directly with youth books. Because only students with low reading levels participated, dyads consisted exclusively of weak readers, eliminating the performance heterogeneity typical of PT programs.

Treatment fidelity was ensured by monitoring adherence to baseline and intervention protocols; interrater reliability was 98% (Sanetti et al., 2021). At the end of the study, social validity was assessed. Students completed a questionnaire (Likert scale 0–4) on perceived benefit, motivation, and enjoyment. Teachers provided feedback on implementation, transfer to regular instruction, and student motivation.

Results

Findings indicated that all participants improved in reading comprehension from baseline to intervention, though to varying degrees. Hakim achieved the largest gain (62.36%), whereas Adem showed the smallest improvement (6.71%). One student, Azad, initially demonstrated a performance decline of 4.52% during the intervention phase but showed substantial improvement in the follow-up phase. Overall, follow-up results were mixed, with some students making further progress while others regressed.

In most cases, the highest performance levels were achieved during the intervention phase (B). An exception was Emre, whose best score occurred during the follow-up phase. Statistical analyses using Tau-U (Parker et al., 2011b) and Non-Overlap of All Pairs (NAP; Parker et al., 2011a) indicated predominantly small to moderate effects: five students showed moderate effects, three showed small effects.

Social validity ratings from students were overall positive. Peer collaboration, the point system, willingness to participate, and reading support were rated most favorably. Lower ratings were given to understanding the program's overall goal, interest in the selected books, and the impact of the points system on concentration. However, one student's low ratings disproportionately reduced mean scores.

Teachers also evaluated the program positively, noting increased student motivation, enthusiasm for the intervention, and application of the strategies in regular instruction. Overall, teachers confirmed the program's usefulness in enhancing reading competence.

Discussion

The findings indicate that the adapted German version of PALS can improve reading comprehension among L2 learners in lower secondary school, although effect sizes varied considerably as a function of vocabulary knowledge and cognitive load. These results are consistent with Snow's conceptual model (2002), which emphasizes the interaction between text, reader, and activity. In this study, progress depended on the use of linguistically accessible, motivating texts and on the successful alternation of roles within PT. The findings also align with research on L2 competence, vocabulary-WM relationships, and the benefits of targeted interventions for students entering secondary school (Shin, 2020).

Social validity results further underscored the program's positive impact: students perceived PALS as motivating and supportive, particularly valuing the four structured steps, which fostered both social skills and reading motivation.

Several limitations must be acknowledged. High attrition was observed: of 24 students initially supported, only eight met the inclusion criteria due to frequent absences. The short six-week intervention period limits conclusions regarding long-term effects on comprehension and fluency. Moreover, the exclusive pairing of weak readers departed from the original PALS design, which relies on heterogeneous dyads. The selected books also proved too linguistically demanding for some participants. Finally, organizational challenges such as scheduling conflicts and absences further reduced sample size.

Despite these limitations, the study suggests that PALS represents a feasible and effective approach for promoting reading competence among L2 learners in German lower secondary schools. With appropriate adaptations, it can be integrated into regular classroom practice. Future research should employ a randomized controlled trial (RCT) to confirm these effects. In addition, tailoring book selection more closely to students' reading levels and interests is recommended. A clearer distinction between L2 acquisition difficulties and genuine learning disorders is also crucial, as language deficits are often prematurely equated with general learning problems. Finally, heterogeneous pairing of weaker and stronger readers may enhance PT dynamics and yield additional gains.

2.4.3 Summary Article 3 *(peer-reviewed)*

Barwasser, A., Bracht, J., Lenz, B., Gürçay, I., Hoff, S., & Grünke, M. (2025) Effects a peer-tutorial story-map intervention on the reading and writing of students with and without German as a second language. *Reading & Writing*, *38*, 1337–1357. https://doi.org/10.1007/s11145-024-10565-0

Building on evidence that PT in secondary school enhances reading—especially for learners of GL2, as shown in the second PALS study—this investigation examined whether PT elements can function as an earlier intervention at the end of primary school. Because reading and writing are tightly coupled, and this link is supported by several meta-analyses (Graham et al., 2017; Graham & Hebert, 2011), this study additionally examined whether the intervention could be combined with structured writing practice. An early start in elementary school is

warranted to prevent later L2 acquisition difficulties that tend to solidify in secondary grades (Barwasser et al., 2025; Valcárcel Jiménez et al., 2024).

Method

We implemented an experimental pre-post-follow-up design with two groups (intervention vs. control) and randomized allocation at the classroom (cluster) level. Participants were 60 children from third- and fourth-grade classrooms. The intervention lasted four weeks (three 45-minute sessions per week; 12 sessions total). All children met risk criteria for reading difficulties (ELFE-II percentile < 30; Lenhard et al., 2020), indicating either manifest problems or elevated risk. 38 children were L2 learners of German, typically since age three. Three children had special needs.

Sentence- and text-level comprehension were assessed with the standardized ELFE-II (Lenhard et al., 2020). Writing quality was rated using a rubric adapted from the Teacher Evaluation of Story Elements scale by Troia and Graham (2002) across five dimensions (setting, problem, actions, consequences, emotions); this instrument was not standardized.

The program combined (a) a story map graphic organizer, (b) the SRSD model with its six phases (develop background knowledge, discuss, model, memorize, support, independent performance), (c) PT elements, and (d) positive reinforcement (stickers, sweets, end-of-program certificate).

Materials comprised a reading folder with 12 stories and a corresponding story map for each. Instruction followed the Gradual Release of Responsibility (GRR) model by Pearson and Gallagher (1983) and was delivered by graduate students in special education. The control group received regular classroom instruction without supplemental reading or writing support.

Treatment fidelity was documented by two external observers per group; interrater reliability was 100% (Sanetti et al., 2021), indicating high adherence. Social validity was assessed via child ratings and teacher feedback.

Results

The intervention group showed significant gains in reading comprehension. For sentence comprehension, posttest: $F(1,56) = 11.74$, $p < .01$, $\eta^2 = .17$; gains remained at two-month follow-up: $F(1,56) = 8.15$, $p < .01$, $\eta^2 = .13$. For text comprehension, effects were larger: posttest $F(1,56) = 32.06$, $p < .001$, $\eta^2 = .36$; follow-up $F(1,56) = 12.87$, $p < .001$, $\eta^2 = .19$.

Mean rubric scores for writing quality increased descriptively from 6.45 to 9.14. A significant posttest difference emerged, $F(1,56) = 3.82$, $p < .05$, $\eta^2 = .06$; however, violation of the homogeneity-of-regression-slopes assumption limits interpretability. No significant differences were observed at follow-up, $F(1,56) = 1.31$, $p > .05$. Children rated the reading folder, partner

work, and rewards as highly motivating. Teachers reported uniformly positive impressions, citing strong acceptability and perceived benefits for both comprehension and writing, and expressed interest in embedding the program into regular instruction.

Discussion

The story map intervention, grounded in SRSD and PT, produced significant improvements in sentence- and text-level comprehension, with the largest effects for text comprehension ($\eta^2 = .36$). Descriptive gains in writing quality suggest potential for supporting text production, though conclusions are constrained by measurement limitations. Notably, the intervention was effective for L2 learners without L2-specific adaptations, and as early as grade three and four.

These findings align with prior research on multimodal, strategy-based interventions leveraging SRSD, PT, and graphic organizers (Sencibaugh, 2007; Sun et al., 2021). As in related work, combining reading and writing strategies proved pedagogically coherent, and explicit teaching of text structure via story maps was especially effective for students with learning difficulties (Cure et al., 2020). The intervention is readily implementable in inclusive, general-education settings and can be flexibly calibrated to students' reading proficiency. Effective delivery requires careful pairing in peer dyads, selection of linguistically accessible and engaging stories, and deliberate scaffolding across the different GRR phases—features embedded in this study's design.

Generalizability is restricted by the modest sample size and the absence of an active-control comparison, precluding differential efficacy claims against alternative programs. Potential moderators—reading fluency and vocabulary—were not incorporated despite their likely influence (De Jong et al., 2012; Lervåg & Aukrust, 2010). Writing was evaluated with a non-standardized rubric, limiting validity. Replication with larger samples and head-to-head comparisons against established reading programs are warranted. Studies should include fluency and vocabulary as moderators, examine long-term effects on writing and transfer, and further validate the social-validity scale.

A story map- and SRSD-based, peer-tutored intervention constitutes an effective early method for enhancing reading comprehension in third- and fourth-grade students with and without L2 backgrounds, with promising—though not yet definitive—benefits for writing. High social acceptability among students and teachers underscores its practical value for inclusive primary classrooms and for supporting early L2 literacy, intervening before difficulties become entrenched.

2.5 Interim Conclusion

The theoretical concepts of literacy presented in this chapter form the foundation for effective reading and writing instruction. The subsequent empirical studies build on these aspects and provide important insights into both the adaptation of existing programs and methods to the German-speaking context and its school structures, and the targeted support of L2 learners—a group that has so far received too little attention in research.

The first two studies on the adapted German version of the PALS program demonstrate that even a translated and abridged version of the original program is capable of confirming international findings on the effectiveness of PALS (Fuchs et al., 1997, 2020, 2021; Lee et al., 2023; Sáenz et al., 2005). Significant improvements in reading comprehension and fluency were documented, including among students with GL2 and those with SWRD. Particularly important were the design of the materials and the integration of motivational reinforcers, both of which played a decisive role in strengthening the effectiveness of the intervention. At the same time, the results showed that younger children, many of whom exhibit substantial vocabulary deficits, require a very close alignment of materials and methods in order to benefit to the same extent as older students.

The story map intervention described in the third study also made a valuable contribution to the existing body of research. By combining different methodological approaches such as PT, SRSD, and graphic organizers, it demonstrated that it is precisely the integration of these components that enables effective support for reading competence and, to some extent, writing competence. These findings build on earlier research and extend it by providing new evidence on the combined effects of these methods for L2 learners in the German-speaking context.

Taken together, all three studies confirmed strategy training and PT as central success factors for literacy support. They proved effective for both younger and older students, provided that the materials and methods were carefully adapted to the learners' prior knowledge and linguistic abilities.

For future research, it will be particularly important to replicate these findings with larger samples, to test the adapted German version of PALS in additional contexts, and to examine further components in a targeted way—for example, the impact on writing competence in the combinatory story map intervention. Collectively, the studies underscore that literacy interventions can be effective across grade levels, yet they must always be closely aligned with the linguistic and cognitive prerequisites of the learners.

3. Early Literacy and Vocabulary Support in German as a Second Language Learners

3.1 Early Literacy Competence in Second Language Learning

To better understand the development of literacy, particularly reading competence, the Dual Route Theory (DRT) provides a valuable cognitive-psychological framework (Coltheart et al., 2001). This model distinguishes between the lexical route and the sublexical (phonological) route in reading.

The lexical route relies on a mental lexicon that enables the rapid and automatic recognition of words, including their orthography, pronunciation, and meaning. This route is crucial for reading fluency. The sublexical route, by contrast, is used to decode unfamiliar words or pseudowords. In this process, PA is applied by synthesizing sounds through grapheme-phoneme correspondence (GPC). Although slower and more effortful, this route is essential for acquiring new words (Castles et al., 2018; Junker et al., 2024).

Both routes operate in parallel, but their relative use shifts with reading experience. At the beginning, the sublexical route dominates, since novice readers cannot yet rely on an established mental lexicon (Hasenäcker & Schroeder, 2022). Every word is initially new, whether in L1 acquisition or in learning a L2. This process follows a self-teaching mechanism (Share, 1995): with every newly decoded word, the child expands their sight-word lexicon until the word can be stored in the mental lexicon after accurate decoding. As reading experience grows, reliance on the sublexical route decreases, while the lexical route becomes increasingly efficient. Continuous exposure to written language is necessary for expanding the mental lexicon (Castles et al., 2018).

For L2 learners, an analogous process applies: they must also build a mental lexicon in the target language and initially depend on the phonological route. Vocabulary plays a central role here, since a decoded word can only be permanently stored in the mental lexicon if its meaning is known (Castles et al., 2018). Contexts or similarities to the L1 may provide important support (Bosma & Nota, 2020).

PA is particularly significant for the sublexical route. It refers to the ability to perceive and manipulate the sound structure of a language (Hasenäcker & Schroeder, 2022). Researchers differentiate between awareness of large phonological units (e.g., recognizing syllables and rhymes) and small phonological units (identifying phonemes and graphemes) (Haase & Steinbrink, 2022). The latter typically develops alongside literacy acquisition and is regarded as a

key predictor of literacy competence. Studies have shown that deficits in PA at the end of pre-school strongly correlate with later reading and spelling difficulties (Valcárcel Jiménez et al., 2024).

In L2 acquisition, the extent to which phonological structures in L1 and L2 overlap is critical, as is the orthographic transparency of the target language (Borleffs et al., 2019). In transparent orthographies (e.g., Italian, Spanish), graphemes usually correspond to single pho-nemes. In deeper orthographies (e.g., English), one grapheme may correspond to multiple pho-nemes. German is considered largely transparent, though it also includes multiple correspond-ences, such as with the letter <v> (*Vater* vs. *Vase*).

GPC forms the foundation of decoding words via the sublexical route. A solid grasp of GPC facilitates quick recognition of letter-sound associations and supports the transition to di-rect word recognition (Hasenäcker & Schroeder, 2022). This is often a challenge for L2 learn-ers, since GPC rules in the target language may differ from those of the L1. Such differences can lead to interference, mispronunciations, and comprehension problems. Transfer is espe-cially difficult for learners whose L1 uses non-alphabetic writing systems (e.g., Chinese) or alphabets with different graphemic structures (e.g., Arabic) (Verhoeven & Perfetti, 2021).

GPC is equally essential for writing competence, though in the reverse direction: sounds must be encoded into written form by mapping phonemes onto graphemes (Breining & Rapp, 2019). The concept of Letter-Sound Fluency (LSF) captures the degree to which GPC has been automatized. It is typically assessed by measuring how many letters a learner can correctly produce sounds for within a given time (Clayton et al., 2020). While PA refers to knowledge and perception of sound structures, LSF reflects the speed and automaticity with which these associations are applied.

Research highlights the predictive value of accuracy and speed in letter-sound recogni-tion across all levels of the reading process. For example, Good et al. (2001) emphasize their importance for later literacy, and Speece and Case (2001) reported a strong correlation ($r = .66$) between LSF at the beginning of first grade and reading performance at the end of the same school year.

The more robust the GPC, the higher the level of LSF, and the stronger the foundation for literacy development. A well-developed GPC strengthens the sublexical route, supports the growth of the lexical route, and thereby contributes to the expansion of the mental lexicon (Cas-tles et al., 2018). Because the permanent storage of words also requires semantic knowledge, vocabulary development plays an equally vital role. This aspect will be examined more closely in the following chapter, with a focus on L2 learners.

3.2 Vocabulary in Second Language Learning

L2 learners frequently display marked vocabulary deficits, both in the early and intermediate stages of L2 acquisition. Numerous studies have shown that their vocabulary size lags significantly behind that of L1 learners (Babayiğit et al., 2022; Hjetland et al., 2023). The German English Student Performance International (DESI) study (Klieme, 2006), for example, demonstrated that ninth-grade students with GL2 had considerable vocabulary gaps compared to their monolingual peers. These differences are already evident in elementary school: children with GL2 score significantly lower in receptive vocabulary than those with German as a L1 (Röthlisberger et al., 2023). Particularly at risk are children who have late exposure to German or do not attend preschool; they face a heightened risk of more pronounced deficits (Sachverständigenrat deutscher Stiftungen für Integration und Migration [SVR], 2013; Statistisches Bundesamt, 2024b).

These gaps often persist even after several years. Without targeted intervention, their negative impact on overall language development and literacy competence in the L2 can hardly be compensated (Nation, 2001). The relationship is clear: vocabulary size correlates strongly with both reading competence and general language proficiency (Lervåg & Aukrust, 2010; Wise et al., 2007). A limited vocabulary further disadvantages learners by making it difficult to follow classroom instruction and understand assignments, which in turn leads to learning difficulties across subjects. Laufer (1989) argues that approximately 95% lexical coverage is necessary for adequate text comprehension. If this threshold is not reached, significant comprehension gaps emerge—a central challenge in L2 acquisition.

Of particular importance are so-called high-utility words, meaning high-frequency words or central function words (e.g., articles, prepositions). In German, as in other languages, a relatively small subset of words accounts for the majority of language use. For English, Nation (2001) identified roughly 2,000 words that cover about 80% of the vocabulary in texts. While this level of coverage alone does not guarantee full reading comprehension, high-utility words are indispensable for basic communication. In academic contexts, learners also require low-utility words, though only once a stable foundation of high-utility words is in place.

Vocabulary is commonly divided into receptive and expressive vocabulary (Hjetland et al., 2023). Receptive vocabulary encompasses words learners understand, while expressive vocabulary refers to words they actively use in speaking or writing. Receptive knowledge is particularly critical for listening and reading, whereas expressive knowledge is essential for speaking and writing. High-utility words are generally acquired through repeated oral and written encounters on both levels, while low-utility words initially remain largely receptive (Lee,

2025). In general, receptive vocabulary is always larger than expressive vocabulary, since active use requires more detailed knowledge, such as correct pronunciation, spelling, and contextual appropriateness (Gilyuk et al., 2021; Heidari, 2024; Lee, 2025). However, the more frequently a word is needed and used, the more likely it is to shift into the expressive vocabulary (Heidari, 2024; Lee, 2025).

Researchers further distinguish between vocabulary breadth and vocabulary depth (Heidari, 2024). Vocabulary breadth refers to the quantity of known words, while vocabulary depth describes the quality of word knowledge. Depth includes understanding polysemy (e.g., *Schloss* as *castle* vs. *lock*), idiomatic uses (e.g., *Schlüssel* in *Schlüssel zum Erfolg* [key to success]), and contextual nuances. Whereas vocabulary breadth tends to grow relatively quickly, vocabulary depth develops more slowly and requires sustained, intensive exposure to the language (Heidari, 2024; McKeown, 2019). A well-established core vocabulary of high-utility words is therefore essential for increasing vocabulary breadth. At the same time, the development of vocabulary depth is necessary to enable precise and flexible language use, which is crucial for advancing literacy skills.

Vocabulary-related hurdles are multifaceted and must not be underestimated, as they significantly contribute to weaker literacy outcomes (Dixon et al., 2022; Karipidis et al., 2018). Within the framework of the DRT, this means that solid PA must be combined with the development of a growing vocabulary—particularly in high-utility words—in order to support literacy development effectively. Difficulties in one or more of these areas should therefore be identified as early as possible and addressed through targeted intervention.

3.3 Enhancing Vocabulary and Early Literacy Competence

For the development of strong early literacy skills and a broad vocabulary in a L2, appropriate learning methods are essential to address barriers in L2 acquisition before they become entrenched. The preceding chapters have shown how strongly literacy competence in the early years is predicted by well-developed early literacy skills (e.g., PA) and, above all, by a broad foundational vocabulary. The learning of these skills can be divided into two major domains: incidental learning (implicitly acquired knowledge) and intentional learning (consciously and explicitly taught knowledge) (Webb & Nation, 2017).

In incidental learning, students acquire skills more or less "along the way," for example, by learning new words through intensive language contact and an environment in which language is lived. Storytelling is a particularly effective method: a narrator reads a story aloud,

which is followed together, and uses gestures and facial expressions to make the story engaging (Barwasser et al., 2021, 2022; Hostetter, 2011; Isbell et al., 2004). This can be complemented by dialogic reading, in which learners are drawn into an exchange about the text through questions, comments, or repetitions (Ceyhan & Yıldız, 2021). A range of studies confirm the effectiveness of storytelling and dialogic reading (Barwasser et al., 2021, 2022; Ceyhan & Yıldız, 2021; Isbell et al., 2004). By promoting active language use and meaningful, context-bound encounters with language, learning can take place outside of traditional classroom structures. Younger children in particular benefit from this method, as it helps them compensate for early vocabulary gaps. At the same time, shared reading and expressive read-alouds foster reading enjoyment and a positive attitude toward books and stories (Ceyhan & Yıldız, 2021).

The central strength of incidental learning lies in context: words are not taught in isolation but embedded in meaningful discourse. Learners acquire far more than the simple decoding of a word—they experience directly how it is used semantically and grammatically in sentences, which in turn enhances reading comprehension. This is especially effective when target vocabulary is used richly and repeatedly within a text, allowing learners to encounter words in varied contexts that deepen their understanding (Uchihara et al., 2019). In this way, both vocabulary breadth and vocabulary depth are strengthened. Incidental learning thus creates natural language situations and promotes meaningful, contextualized acquisition. L2 learners benefit greatly, as it allows them to develop pronunciation, literacy, and vocabulary depth simultaneously (Nation, 2013; Neumann et al., 2022; Uchihara et al., 2019).

Intentional learning, by contrast, refers to explicit and targeted instruction in a learning item such as vocabulary. Unlike incidental learning, the learning situation is clearly structured and consciously recognized as such. A typical method is direct instruction (DI): target words are deliberately selected and systematically introduced and practiced with students, with age-appropriate explanations and less reliance on implicit or "hidden" creativity than in incidental learning (Gallagher et al., 2019). Meta-analyses confirm the effectiveness of intentional learning. Marulis and Neuman (2010), for example, found that direct vocabulary instruction was more effective than exposure to language alone through implicit learning. When students learn words deliberately, retention is stronger than when they encounter them incidentally; however, transfer effects remain limited, so varied usage contexts are still crucial in intentional learning.

Intentional learning is well complemented by Paivio's Dual Coding Theory (DCT; 1991), which posits that information is better retained when presented both verbally and visually. This principle can be easily applied in DI through flashcards or other visualizations: whenever a new learning item or word is introduced and practiced, pictures, illustrative gestures, real

objects, or other forms of visualization support memorization and provide a multisensory learning experience (Mathias et al., 2021). Attention is directed simultaneously to auditory and visual input—an approach especially helpful for L2 learners, who often rely heavily on visual anchors (Mathias et al., 2021).

Motivation is a central factor, particularly for younger students in early literacy development. Visual materials can be tailored to their interests; linguistic imagery or objects that evoke an emotional response (for example, humorous elements) are more likely to be remembered (Hall et al., 2021). A connection to children's everyday lives—for instance, familiar themes or characters—also sustains motivation. Animal stories, for example, often spark natural interest and maintain attention (Flynn, 2004). Such connections to lived experience create a sense of relevance that children readily perceive.

Making progress visible is another powerful motivator. Self-graphing, as part of self-monitoring, is one effective method (Guzman et al., 2018): words that have been deliberately practiced are tracked in tasks or assessments and displayed as a performance curve. The more words mastered, the steeper the curve rises—making success immediately visible and traceable over time. However, this approach carries risks: if progress plateaus, it may become demotivating. In such cases, praise and positive reinforcement are more important than the strict visualization of performance. Positive reinforcement has been shown to foster learning, strengthen positive self-perceptions, and encourage persistence (McLeskey et al., 2017). Tangible rewards (such as stickers or small tokens) can also signal that effort is valued and achievements recognized.

Across all methods, the principle holds: the earlier the better. Because early literacy precursors are already critical in elementary school, barriers in these areas must be addressed as soon as possible. Hart and Risley's (1995) "30-million-word gap" highlights the urgency: by age three, children from language-poor families have heard approximately 30 million fewer words than peers from linguistically stimulating environments. Preschool therefore plays a pivotal role in creating an enriched linguistic atmosphere (SVR, 2013; Statistisches Bundesamt, 2024b)—fertile ground for the development of literacy competence through diverse and complementary approaches.

3.4 Publications

The preceding considerations have highlighted which linguistic precursor skills are essential for successful literacy development, particularly PA and vocabulary. The DRT emphasizes that GPC interacts closely with a continuously expanding mental lexicon, and together they form the foundation of solid literacy competence.

For L2 learners, additional hurdles arise at precisely this point. They often face difficulties not only with GPC in the L2 but also with limited vocabulary, which significantly hinders literacy acquisition (Karipidis et al., 2018). Instructional approaches that combine incidental and intentional learning can help address these challenges. Incidental methods such as storytelling provide contextualized and natural encounters with language, while intentional methods such as DI explicitly and systematically target linguistic gaps (Law et al., 2019). It is also crucial to incorporate motivation and visual support in order to foster positive learning experiences and create linguistically stimulating environments (Feng & Webb, 2020; Karipidis et al., 2018).

The following two studies take up these aspects and examine how storytelling interventions, both with and without DI, influence the development of vocabulary, GPC, and early reading skills in children learning GL2. The focus is deliberately placed on young learners at the beginning of their schooling and even in the preschool phase, in order to address linguistic barriers as early as possible. In doing so, these studies address a central research gap by systematically investigating early childhood interventions that integrate elements of incidental and intentional learning within the German-speaking context for L2 learners.

3.4.1 Summary Article 4 *(peer-reviewed)*

Barwasser, A., Bracht, J., & Grünke, M. (2021). A storytelling approach on vocabulary, reading and letter sound fluency of struggling first graders with German as second language with and without behavioral problems. *Frontiers in Psychology*, *12*, article 683873. https://doi.org/10.3389/fpsyg.2021.683873

Since precursor skills for reading are of critical importance, it is essential to foster them at an early stage in order to reduce difficulties in L2 acquisition and thereby prevent the emergence of later reading problems (Valcárcel Jiménez et al., 2024). Against this background, we conducted a study to examine the impact of a storytelling intervention on expressive vocabulary, LSF, and sight-word reading.

Method

The study employed a multiple-baseline across participants design with a total of seven children who took part in the intervention. The participants were first graders acquiring GL2. This constituted a very early intervention, implemented directly at the beginning of formal schooling, and had the potential to act as a preventive measure against subsequent reading difficulties. The children's home languages were diverse and included Turkish, Polish, Chinese, and Italian. Some of the children exhibited behavioral difficulties, which were assessed using the ITRF-G teacher questionnaire (Volpe et al., 2018). The inclusion criterion for participation in the study was a percentile rank below 15 on reading and language tests. The following assessments were administered: the Vocabulary Test and Number Sequence Test Revision (WS/ZF-R; Weiß, 2007) for expressive vocabulary, the further development of the Salzburg Reading and Spelling Test (SLRT-II) for word and pseudoword reading, the Basic Skills for Reading and Spelling Test (BAKO 1–4) for PA, as well as researcher-developed tests for 40 training words and LSF, in which all letters were tested.

The intervention lasted approximately six weeks and consisted of three sessions per week, each lasting 25 minutes. Each session was divided into two phases. The first phase involved explicit training with DI, in which the children practiced target words and letters using flashcards and a phoneme chart. This was followed by the second phase, a storytelling component in which researcher-developed stories embedded the target words. To enhance motivation, the children used self-graphing sheets to visually track their progress.

The dependent variables were expressive vocabulary, measured as the number of correctly named target pictures, LSF, measured as the number of correctly named sounds per minute, and sight-word reading, measured as the number of correctly read words with one second allowed per word. Data analysis was carried out through both visual inspection and statistical procedures, including Tau-U, NAP, Percentage of Data Points Exceeding the Median (PEM; Ma, 2006), and Percentage of All Non-Overlapping Data (PAND; Parker et al., 2007). In addition, a hierarchical regression analysis using the Scan package in R (Wilbert & Lueke, 2021) was performed to examine level and trend changes from the baseline (A phase) to the intervention (B phase).

Results

In expressive vocabulary, all children achieved significant and in some cases very large gains. The greatest improvements were observed for Kim with an increase of 3,034% and Lio with 2,469%. The highest mean scores in the B phase were achieved by Tila ($M = 31.42$) and Elif ($M = 30.58$). All Tau-U values were significant and indicated large to very large effects.

The regression analysis confirmed these findings through significant increases during the B phase.

For LSF, baseline performance was heterogeneous, and some children already showed positive trends in the A phase. Nevertheless, all children improved, with particularly strong gains for Abden (116.20%) and Tila (112.59%). The Tau-U values were again significant and indicated moderate to large effects. The regression analysis revealed significant level changes in only one of the three groups ($p < .05$).

Sight-word reading emerged as the weakest domain, with only gradual improvements. Only Nele reached the maximum of 40 words, while the other children scored considerably lower. The largest improvements were recorded for Elif with 2,922,39% and Nele with 2,562%. Both children also showed very large Tau-U effects, and regression analyses indicated significant increases in two of the three intervention groups.

Social validity was also assessed in order to capture participants' acceptance of the intervention. The results revealed very high levels of agreement. All children reported that storytelling helped them to read, speak, and understand better. The use of self-graphing was also consistently rated positively, as it promoted motivation and metacognition.

Discussion

In summary, the multicomponent storytelling intervention with DI demonstrated very strong effects. The combination of storytelling and explicit instruction proved to be the central success factor. Particularly in the domain of vocabulary, all children achieved substantial training gains. These gains were further influenced by prior experiences, attention, and learning behavior, findings that have also been reported in earlier studies (e.g., Barwasser et al., 2021; Marulis & Neumann, 2010).

LSF also improved significantly. Since this skill is considered an important predictor of later reading fluency (Hulme et al., 2012), these findings are especially relevant. Despite substantial differences at baseline, even children with low PA made marked progress. Sight-word reading showed the least improvement, which can be explained by the longer process of automatization required (De Jong et al., 2012). This process is described in the DRT, which incorporates both the lexical and the non-lexical route. It can also be assumed that PA and L1 exerted an influence on the outcomes.

Self-graphing further enhanced motivation and metacognition (Guzman et al., 2018) and proved to be an important factor for sustained learning. Taken together, the findings indicate that the intervention can be easily integrated into daily school practice and is particularly

effective for children acquiring GL2 already in early elementary school. Thus, this multicomponent intervention represents a promising approach for the prevention of later reading difficulties.

3.4.2 Summary Article 5 *(peer-reviewed)*

Bracht, J., Wasko, L., Grünke, M., & Barwasser, A. (2025_accepted). Storytelling with and without direct instruction on grapheme-phoneme correspondence and vocabulary of preschool children learning German as a second language. *Learning Disabilities: A Contemporary Journal*.

Early intervention is of particular importance for children at risk for reading difficulties or for those who already exhibit initial indicators of impaired development—such as deficits in PA or a limited sight-word vocabulary. For these children, systematic support targeting precursor skills at an early stage is critical in order to establish a stable foundation for reading competence and to ensure age-appropriate progress during elementary school relative to their peers.

Intervention must therefore begin prior to formal schooling. In the preschool years, the primary emphasis should be on precursor skills relevant to reading acquisition. In the German educational system, however, neither a standardized national preschool curriculum nor compulsory attendance exists. Preschool typically corresponds to the final year of kindergarten before school entry. Given that preschool attendance and structured intervention are not mandatory, opportunities for early literacy experiences are highly unequal. For some children, initial exposure to literacy occurs only upon school entry. Yet difficulties manifesting in elementary school often become rapidly entrenched and persist across subsequent grades. Early intervention has the potential to counteract this trajectory. Children acquiring GL2 particularly benefit from early and intensive language exposure, which facilitates acquisition of the phonological system and vocabulary of the L2 (Lesaux et al., 2007; Valcárcel Jiménez et al., 2024). To address these issues, we conducted a preschool-based study investigating whether intervention in PA, GPC, and expressive vocabulary yields measurable short-term gains among children at heightened risk, including those acquiring GL2.

Method

The study employed a single-case multiple-baseline across participants consisting of three phases: A (baseline), B (incidental storytelling), and BC (storytelling combined with DI). In the B phase, children were presented with a story embedding ten target words. These items

were integrated contextually through gesture, facial expression, and hand puppets without explicit instruction. In the BC phase, explicit instruction was added: target words were introduced with flashcards, discussed, phonologically segmented, and repeated by the children. Thus, incidental learning was supplemented with systematic, explicit practice.

Participants were seven children aged five to six years with migration backgrounds. Their L1s included Turkish, Russian, Arabic, and Kurdish. The intervention was implemented over eight weeks in a kindergarten in North Rhine-Westphalia, Germany. To ensure staggered baselines and strengthen internal validity (Kratochwill et al., 2013), children were divided into three groups. Each child completed 21 measurement sessions, with intervention phase durations varying across groups.

Assessment instruments included the Test for the Acquisition of Phonological Awareness and Naming Speed (TEPHOBE; Mayer, 2020) for PA and rapid automatized naming, a LSF task (number of correctly identified sounds within 60 seconds), a GPC task (PowerPoint-based letter identification), and an expressive vocabulary task (naming of 30 pictured items presented via PowerPoint). GPC and expressive vocabulary were assessed repeatedly at all measurement points. Social validity was additionally evaluated at the study's conclusion: children completed visually supported questionnaires, and educators rated the intervention on a Likert scale.

Results

Findings indicated negligible gains in expressive vocabulary between phases A and B; five children maintained scores of zero throughout this interval. Significant improvements emerged only in the BC phase, where storytelling was combined with explicit instruction. All children demonstrated robust gains. One participant, Defne, achieved near ceiling performance (29/30 correct). Group means ranged from 2.13 to 15.25, substantially exceeding baseline scores and documenting clear improvement. Effect size analyses (Tau-U, Percentage of Data Exceeding a Median Trend [PEM-T; Alresheed et al., 2013; Ma, 2006]) corroborated these results: no significant differences were observed between A and B, but large effects emerged in the BC phase across all children (Tau-U = .73–.97, $p < .001$).

Similar patterns were observed for GPC. Performance remained stable across phases A and B, with minimal improvement. Substantial gains were observed only in the BC phase, particularly among children with low initial performance. For example, Defne's mean score increased from 1.0 to 10.18, with a maximum of 17 correct responses. Alara improved from a mean of 9.80 to 17.13, achieving up to 20 correct responses. Aras, previously scoring at floor, achieved a mean of 2.13. Effect size calculations confirmed these outcomes: while no

significant effects emerged between A and B, significant effects were documented for six of seven children in the BC phase (Tau-U values up to .94, *p* < .001).

Social validity outcomes were largely positive. Five of seven children consistently provided maximal ratings ("green thumb"), while two were absent at the final session. Educators assessed the intervention favorably, highlighting the motivational value of the storytelling format. Concerns were raised, however, regarding program scope and the limited visibility of immediate effects in daily practice, complicating attribution of observed progress. Nonetheless, most educators indicated willingness to integrate the intervention into their pedagogical repertoire.

Discussion

In summary, storytelling alone did not produce measurable learning gains. Only when combined with DI did significant improvements occur. Children with initially low performance benefitted most, consistent with the so-called "Robin Hood effect" (Häfner et al., 2017). From a practical perspective, these findings suggest that weaker learners can accelerate their progress and narrow achievement gaps when implicit learning (storytelling) is complemented by explicit instruction (flashcards). The results also align with DCT, as linking verbal input with visual and auditory encoding facilitated the intended learning outcomes (Paivio, 1991).

Several limitations warrant consideration. The heterogeneity of children's language backgrounds was not systematically documented (e.g., degree of L1 dominance, extent of L2 exposure). Moreover, no follow-up phase was included, preventing conclusions about long-term effects. This limitation primarily reflected structural constraints, as children transitioned to different elementary schools following preschool. An additional C phase would also have been informative in isolating the effects of DI independent of storytelling. Finally, potential reproduction effects cannot be excluded, as training words remained constant across sessions, although their sequence was randomized.

From a practical standpoint, the findings underscore the value of early, systematic integration of interventions combining incidental and intentional learning into preschool education. The combination of storytelling and DI proved both effective and feasible for everyday practice. Children with language delays or elevated risk for LD benefitted disproportionately. Future research should investigate long-term effects as well as the isolated efficacy of DI.

3.5 Interim Conclusion

Without targeted support that is adapted to learners' needs, deficits can become entrenched at an early stage and, over time, develop into significant barriers to literacy acquisition. The two studies have shown that storytelling, as a motivating and context-based method of incidental learning, provides a valuable foundation for early language instruction. It contributes to the development of GPC, LSF, and vocabulary—key building blocks for later reading and writing skills.

At the same time, the findings demonstrate that sustainable learning progress is achieved primarily when storytelling is combined with DI as part of intentional learning. Significant improvements were observed both in preschool and in elementary school. These results confirm earlier findings on the interplay between incidental and intentional learning and highlight that combining both approaches is essential for making learning both engaging and lasting (Gallagher et al., 2019). Storytelling proves to be a particularly effective method for embedding children early on in linguistically rich environments while also fostering enthusiasm for texts and reading.

The results extend previous research on storytelling and, in particular, provide evidence of its effectiveness and necessary adaptation for L2 learners in the German-speaking context (Barwasser et al., 2021; 2022). They make clear that this group of students benefits substantially from early interventions. In the long term, such measures can effectively reduce literacy difficulties that stem either from limited L2 skills or from LD.

It should be noted, however, that these studies are based on single-case-research designs (SCR) with small samples and relatively short intervention periods. Long-term effects of interventions beginning this early would therefore be of particular research interest. Nevertheless, the findings underscore the considerable potential of these approaches for the German-speaking context and provide compelling evidence that interventions can and should begin very early in L2 acquisition. The earlier problems are identified and purposefully addressed, the lower the risk that they will develop into serious barriers later in the educational trajectory.

4. Overall Conclusion and Implications

4.1 Conclusion

This dissertation examined, across five studies, the extent to which adaptive interventions at different age levels can address the heterogeneous linguistic demands and needs of children and adolescents learning GL2, and how such interventions can support the development of literacy competence in a comprehensive sense. For older students, the focus lies primarily on reading and writing competence, while for younger children the emphasis is placed on PA and vocabulary. The aim is to provide low-threshold, practice-oriented methods from preschool through adolescence that strengthen learners' abilities, foster motivation, and counteract linguistic barriers as early as possible by targeting support precisely where it is most urgently needed. Gaps in language competence among L2 learners often become apparent only when the gap between required and actual skills grows so wide that it appears nearly insurmountable, particularly in reading or writing extended texts.

The first study investigated the effectiveness of an adapted PALS (Fuchs et al., 1997) intervention in seventh grade among students who already exhibited manifest reading difficulties or who were at high risk of developing them, including L2 learners. The goal was to improve both reading comprehension and fluency. Findings from this pilot study showed that even after a short intervention period, significant gains were observed, which remained stable in the follow-up phase. This demonstrates that a program originally developed for L1 learners in the Anglo-American context can also prove effective in the German-speaking setting when carefully adapted to the target group. The high level of structure and methodological clarity were particular strengths of the program, although research suggests that older students may benefit most from strategy training (Wu, 2022).

The second study therefore examined a more extensively adapted form of the PALS program with exclusively GL2 learners in grades five and six. Results indicated that several sessions were necessary to achieve significant effects, yet gains in reading comprehension were evident even after a relatively short period. Both teachers and students evaluated the intervention positively, suggesting that its implementation in regular school practice is feasible.

That peer tutoring can also be applied effectively with younger students was confirmed by the third study, which focused on a story map intervention. Here, third- and fourth-grade students were supported through a combination of SRSD elements, story maps, and PT. Because reading and writing are closely interrelated, the simultaneous promotion of both domains provided clear practical benefits (Graham & Hebert, 2011). Even after a brief intervention, reading

performance improved significantly. Writing also showed notable progress; however, the Analysis of Covariance (ANCOVA) results were not meaningful because the assumption of homogeneity of regression slopes was violated. Social validity data indicated high acceptance and practical feasibility of the intervention. The combination of SRSD, PT, and story maps thus proved promising for supporting literacy even among L2 learners.

The fourth study placed even greater emphasis on the narrative dimension by examining the storytelling method. By integrating facial expressions, gestures, and visual representations, its effects can be amplified in line with Paivio's DCT (1991). At the beginning of formal schooling, when academic demands are new, children require supportive linguistic environments that emphasize incidental learning to facilitate acquisition and foster motivation (Brewster et al., 2002; Wright, 2013). Results revealed large to very large gains in vocabulary and moderate to large effects on reading skills and LSF. Storytelling thus proved to be an effective method for addressing multiple areas of early literacy simultaneously. The vocabulary gains are particularly important, as L2 learners often display significant deficits in this area that strongly impede literacy development (Ender, 2016; Lervåg & Aukrust, 2010). Notably, the intervention also proved effective for children with behavioral difficulties, pointing to the motivating and adaptive nature of the method.

The fifth study focused on preschoolers, comparing storytelling alone as a form of incidental learning with a combined approach of storytelling plus intentional learning through flashcards. Results showed that storytelling by itself did not lead to significant gains. Only when combined with DI were stable and substantial improvements observed in expressive vocabulary and GPC. These findings align with those of Baron and Arbel (2022) as well as the meta-analyses of Marulis and Neuman (2010), which highlight the effectiveness of combining implicit and explicit methods. The study thus provides clear evidence that intervention before school entry can be both valuable and effective in mitigating early language deficits.

In summary, early interventions benefit particularly from the combination of incidental and intentional learning. Storytelling positively impacts vocabulary and PA but must be complemented by direct instruction. Older students benefit from strategy-based programs such as PALS or from integrated approaches like story maps combined with PT and SRSD, which support both reading and writing competence. Across all five studies, adaptively designed interventions—tailored to learners' age, developmental stage, and linguistic prerequisites—proved effective and feasible for integration into school practice.

The findings underscore that adaptive literacy support can yield significant learning gains at any age, provided it is well-structured, age-appropriate, and responsive to learners'

specific needs. Although longer interventions consistently produce stronger effects (Moeyaert et al., 2021; Tang et al., 2021), the present studies demonstrate that even short, intensive interventions can be effective—a crucial finding in light of the limited time resources available in school contexts.

These results align with existing research showing that L2 learners encounter particular challenges across all stages of education, from limited vocabulary and reduced PA to manifest reading and writing difficulties and diminished self-efficacy (OECD, 2019; 2023). Comparative studies such as PISA and IQB (Institute for Quality Development in Education [IQB], 2022; OECD, 2023) highlight that without targeted support, the achievement gap between L1 and L2 learners widens steadily over the course of schooling. Against this backdrop, the need becomes evident for early, specifically adapted interventions that can be integrated into everyday schooling in the German-speaking context to prevent long-term educational disadvantage and social exclusion (Anger et al., 2024; Grünke & Bracht, 2025).

The studies presented here provide concrete approaches for designing effective support measures and indicate which interventions are most suitable at different developmental stages. From these findings, several implications for both future research and educational practice can be derived, which will be elaborated in the following section.

4.2 Implications and Future Research

The studies presented in this dissertation examined different age groups with particular attention to transitional phases—for example, entry into primary school or the transition to secondary school after four years of primary education. Such phases are always associated with unique challenges: familiar structures dissolve, new demands arise, and expectations for learning change significantly. It is therefore especially important to examine these transitions closely and to understand which measures can help students navigate them successfully.

The studies included in this work presented a range of effective approaches tailored to different developmental phases. Future research should build on these findings by investigating how learning environments can be designed to remain as accessible and engaging as possible while also identifying barriers and addressing them through targeted support. Across the five studies, several implications emerged for both research and practice, emphasizing that scientific findings realize their full potential only when they are implemented in classrooms.

One central outcome was the effectiveness of combined approaches. Strategy training paired with PT, or the integration of incidental and intentional learning, proved particularly

powerful in combination. However, open questions remain as to which components contribute most to the observed effects. Is it the interaction between approaches, or would, for example, DI alone be sufficient while the narrative quality of storytelling plays a merely complementary role? To answer such questions, further evidence-based studies are needed, ideally in the form of RCTs. For instance, it would be valuable to test whether an adapted version of PALS differs in its effects from another PT program, or whether story maps produce similar results even without the peer component.

Existing research suggests that incidental and intentional learning are most effective in combination (Gallagher et al., 2019). Applying this insight directly to the interventions studied here would be an important step toward determining the optimal use and "dosage" of individual methods. For practice, this would mean interventions that are not only more effective but also more resource-efficient in the school context. The story map intervention, in particular, demonstrates that combined approaches can promote multiple aspects of literacy simultaneously, such as vocabulary, LSF, GPC.

PT played a prominent role across all studies. Both research and practice reports point to substantial learning gains through this method (Bowman-Perrott et al., 2013; Moeyaert et al., 2021). Yet open questions remain about the most effective pairing: do students benefit more from heterogeneous pairings (stronger and weaker learners together) or from homogeneous teams? Qualitative investigations would also be valuable to analyze how students support each other, what forms of feedback they provide, and which are most conducive to learning. At the same time, it is crucial to examine how teachers can be prepared for and supported in implementing PT so that the method can reach its full potential.

Motivation and social validity were also central themes across the studies. For future research, this underscores the importance of including qualitative perspectives and employing mixed-methods designs. Only when the perspectives of students and teachers inform intervention development can programs be both effective and feasible. Furthermore, the specific drivers of motivation should be examined in greater detail—for example, by comparing different incentive systems or analyzing the effects of interest-based content such as animal stories or real-world topics (Flynn, 2004).

The adaptivity of support was at the heart of this dissertation's research question: How can the heterogeneous learning needs of L2 learners across all age groups be addressed so that they receive the support they require? Achieving this demands a stronger integration of diagnostics and instruction. Developing specific assessment tools for students with GL2, as well as systematically recording individual language biographies, could allow learning profiles to be

captured more precisely and interventions to be designed with greater accuracy. Moreover, the potential of multilingualism should be emphasized more strongly: rather than focusing primarily on deficits (Grosjean, 1985), existing linguistic resources could be harnessed to overcome literacy barriers more effectively.

Taken together, the findings of this dissertation underscore the need for adaptive and practice-oriented literacy interventions that begin early and remain effective across all age levels. Linguistic barriers should not only become visible once major deficits have already emerged; rather, support must be provided immediately when difficulties are first identified. Rich early language environments can offset many problems that would otherwise solidify during the school years. At the same time, it is equally important not to assume that older students are "too late" for improvement. The studies clearly demonstrate that with the right methods and adaptations, learning gains are possible at any age.

In this way, the dissertation makes a meaningful contribution to the existing research. It shows that adaptive literacy support for the diverse learning needs of L2 learners across different age groups is not only possible but also effective—and this work must be carried forward.

> "Whereof what's past is prologue, what to come
> In yours and my discharge!"
> (William Shakespeare, *The Tempest* 2.1.254f)

Bibliography

Ahmed, Y., Kent, S., Cirino, P. T., & Keller-Margulis, M. (2021). The not-so-simple view of writing in struggling readers/writers. *Reading & Writing Quarterly*, *38*(3), 272–296. https://doi.org/10.1080/10573569.2021.1948374

Allen, L. K., & McNamara, D. S. (2020). Defining deep reading comprehension for diverse readers. In E. B. Moje, P. P. Afflerbach, P. Enciso, & N. K. Lesaux (Eds.), *Handbook of Reading Research* (vol. V, pp. 261–276). Routledge. https://doi.org/10.4324/9781315676302-14

Alresheed, F., Hott, B. L., & Bano, C. (2013). Single subject research: A synthesis of analytic methods. *The Journal of Special Education Apprenticeship*, *2*(1). https://doi.org/10.58729/2167-3454.1015

Anger, C., Betz, J., & Plünnecke, A. (2024). *Bildungsmonitor 2024: Studie im Auftrag der Initiative Neue Soziale Marktwirtschaft (INSM)* [Education Monitor 2024: Study commissioned by the New Social Market Economy Initiative]. INSM.

Babayiğit, S., Hitch, G. J., Kandru-Pothineni, S., Clarke, A., & Warmington, M. (2022). Vocabulary limitations undermine bilingual children's reading comprehension despite bilingual cognitive strengths. *Reading and Writing*, *35*, 1651–1673. https://doi.org/10.1007/s11145-021-10240-8

Bandura, A. (1978). Self-efficacy: Toward a unifying theory of behavioral change. *Advances in Behaviour Research and Therapy*, *1*(4), 139–161. https://doi.org/10.1016/0146-6402(78)90002-4

Baron, L. S., & Arbel, Y. (2022). An implicit-explicit framework for intervention methods in developmental language disorder. *American Journal of Speech-Language Pathology*, *31*(4), 1557–1573. https://doi.org/10.1044/2022_AJSLP-21-00172

Barwasser, A., Lenz, B., & Grünke, M. (2021). A multimodal storytelling intervention for improving the reading and vocabulary skills of struggling German as a second language adolescents with learning and behavioral problems. *Insights into Learning Disabilities*, *18*(1), 29–51.

Barwasser, A., Skrobek, H., & Grünke, M. (2025). Phonological awareness training on recoding of at-risk German kindergarten children with and without behavioral difficulties. *International Journal for Research in Learning Disabilities*, *8*(1), 51–65.

Barwasser, A., Urton, K., Knaak, T., & Grünke, M. (2022). Effects of a multicomponent storytellling intervention on the vocabulary acquisition of German L2 adolescents with

learning disabilities and peer-relation difficulties. *Insights into Learning Disabilities*, *19*(2), 173–192.

Bogaerds-Hazenberg, S. T. M., Evers-Vermeul, J., & van den Bergh, H. (2020). A meta-analysis on the effects of text structure instruction on reading comprehension in the upper elementary grades. *Reading Research Quarterly*, *56*(3), 435–462. https://doi.org/10.1002/rrq.311

Borleffs, E., Maassen, B. A. M., Lyytinen, H., & Zwarts, F. (2019). Cracking the code: The impact of orthographic transparency and morphological-syllabic complexity on reading and developmental dyslexia. *Frontiers in Psychology*, *9*, article 2534. https://doi.org/10.3389/fpsyg.2018.02534

Bosma, E., & Nota, N. (2020). Cognate facilitation in Frisian–Dutch bilingual children's sentence reading: An eye-tracking study. *Journal of Experimental Child Psychology*, *189*, article 104699. https://doi.org/10.1016/j.jecp.2019.104699

Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review*, *42*(1), 39–55. https://doi.org/10.1080/02796015.2013.12087490

Brandenburg, J., Fischbach, A., Labuhn, A. S., Rietz, C. S., Schmid, J., & Hasselhorn, M. (2016). Overidentification of learning disorders among language-minority students. Implications for the standardization of school achievement tests. *Journal for Educational Research Online*, *8*(1), 42–65.

Breining, B., & Rapp, B. (2019). Investigating the mechanisms of written word production: Insights from the written blocked cyclic naming paradigm. *Reading and Writing*, *32*(1), 65–94. https://doi.org/10.1007/s11145-017-9742-4

Brewster, J., Ellis, G., & Girard, D. (2002). *The primary English teacher's guide*. Pearson Education.

Cadierno, T., Hansen, M. B., Lauridsen, J. T., Eskildsen, S. W., Fenyvesi, K., & Hannibal Jensen, S. (2022). A multi-factor study of the development of English receptive skills by young Danish children. *Bilingualism: Language and Cognition*, *25*(4), 566–578. https://doi.org/10.1017/S1366728921001085

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Ceyhan, S., & Yıldız, M. (2021). The effect of interactive reading aloud on student reading comprehension, reading motivation and reading fluency. *International Electronic Journal of Elementary Education*, *13*(4), 421–431. https://doi.org/10.26822/iejee.2021.201

Chen, J., Zhang, L. J., & Parr, J. M. (2021). Improving EFL students' text revision with the self-regulated strategy development (SRSD) model. *Metacognitive Learning*, *17*, 1–21. https://doi.org/10.1007/s11409-021-09280-w

Chen, M.-C., Pustejovsky, J. E., Klingbeil, D. A., & Van Norman, E. R. (2023). Standardized mean differences: Flexible methods for single-case designs. *Journal of School Psychology*, *98*, 16–38. https://doi.org/10.1016/j.jsp.2023.02.002

Clayton, F. J., West, G., Sears, C., Hulme, C., & Lervåg, A. (2020). A longitudinal study of early reading development: Letter-sound knowledge, phoneme awareness and RAN, but not letter-sound integration, predict variations in reading development. *Scientific Studies of Reading*, *24*(2), 91–107. https://doi.org/10.1080/10888438.2019.1622546

Coltheart, M., Rastle, K., Perry. C., Langdon, R., & Ziegler, J. (2001). DRC. A dual route cascades model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Cruz Neri, N., Guill, K., & Retelsdorf, J. (2021). Language in science performance: Do good readers perform better? *European Journal of Psychology of Education*, *36*(1), 45–61. https://doi.org/10.1007/s10212-019-00453-5

Cummins, J. (1979a). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*(2), 222–251. https://doi.org/10.2307/1169960

Cummins, J. (1979b). Cognitive / academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, *19*, 121–129.

Cure, G., Batu, E. S., & Gulboy, E. (2020). Effectiveness of the story-mapping strategy in students with disabilities: Meta-analysis of single-case experimental design studies. *Reading & Writing Quarterly*, *37*(6), 513–534. https://doi.org/10.1080/10573569.2020.1858211

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/10.1080/09297 049.2011.625354

De Vita, F., Schmidt, S., Tinti, C., & Re, A. M. (2021). The role of working memory on writing processes. *Frontiers in Psychology*, *12*, article 738395. https://doi.org/10.3389/fpsyg.2021.738395

Dixon, C., Hessel, A., Smith, N., Nielsen, D., Wesierska, M., & Oxley, E. (2023). Receptive and expressive vocabulary development in children learning English as an additional language: Converging evidence from multiple datasets. *Journal of Child Language*, *50*(3), 610–631. https://doi.org/10.1017/S0305000922000071

Dollmann, J., & Weißmann, M. (2020). The story after immigrants' ambitious educational choices: real improvement or back to square one? *European Sociological Review*, *36*(1), 32–47. https://doi.org/10.1093/esr/jcz042

Dong, Y., & Chow, B. W. Y. (2022). Home literacy environment and English as a second language acquisition: A meta-analysis. *Language Learning and Development*, *18*(4), 485–499. https://doi.org/10.1080/15475441.2021.2003197

Duarte, J., García-Jimenez, E., McMonagle, S., Hansen, A., Gross, B., Szelei, N., & Pinho, A. S. (2020). Research priorities in the field of multilingualism and language education: a cross-national examination. *Journal of Multilingual and Multicultural Development*, *44*(1), 50–64. https://doi.org/10.1080/01434632.2020.1792475

Elleman, A. M. (2017). Examining the impact of inference instruction on the literal and inferential comprehension of skilled and less skilled readers: A meta-analytic review. *Journal of Educational Psychology*, *109*(6), 761–781. https://doi.org/10.1037/edu0000180

Ender, A. (2016). And explicit cognitive processes in incidental vocabulary acquisition. *Applied Linguistics*, *37*(4), 536–560. https://doi.org/10.1093/applin/amu051

Feng, Y., & Webb, S. (2020). Learning vocabulary through reading, listening and viewing. *Studies in Second Language Acquisition*, *42*(3), 499–523. https://doi.org/10.1017/S0272263119000494

Flynn, S. (2004). Animal stories. In P. Hunt (Ed.), *International companion encyclopedia of children's literature* (2nd ed., pp. 418–435). Routledge. https://doi.org/10.4324/9781315015729-31

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

Fuchs, D., Fuchs, L. S., & Abramson, R. (2020). Peer-assisted learning strategies (PALS): A validated classwide program for improving reading and mathematics performance. In

A. L. Reschly, A. J. Pohl, & S. L. Christenson (Eds.), *Student engagement* (pp. 109–120). Springer International Publishing. https://doi.org/10.1007/978-3-030-37285-9_6

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, *34*(1), 174–206. https://doi.org/10.3102/00028312034001174

Fuchs, L. S., Fuchs, D., Seethaler, P. M., Cutting, L. E., & Mancilla-Martinez, J. (2019). Connections between reading comprehension and word-problem solving via oral language comprehension: Implications for comorbid learning disabilities. *New Directions for Child and Adolescent Development*, *2019*(165), 73–90. https://doi.org/10.1002/cad.20288

Galbraith, D., & Baaijen, V. M. (2018). The work of writing: Raiding the inarticulate. *Educational Psychologist*, *53*(4), 238–257. https://doi.org/10.1080/00461520.2018.1505515

Gallagher, M. A., Barber, A. T., Beck, J. S., & Buehl, M. M. (2019). Academic vocabulary: Explicit and incidental instruction for students of diverse language backgrounds. *Reading & Writing Quarterly*, *35*(2), 84–102. https://doi.org/10.1080/10573569.2018.1510796

Gazula, S., McKenna, L., Cooper, S., & Paliadelis, P. (2017). A systematic review of reciprocal peer tutoring within tertiary health profession educational programs. *Health Professions Education*, *3*(2), 64–78. https://doi.org/10.1016/j.hpe.2016.12.001

Gillespie Rouse, A., & Sandoval, A. (2018). Writing interventions for students with learning disabilities: Characteristics of recent research. *Learning Disabilities: A Multidisciplinary Journal*, *23*(2), 1–17. https://doi.org/10.18666/LDMJ-2018-V23-I2-8990

Gilyuk, V., Edmonds, A., & Sneed German, E. (2021). Exploring the evolution in oral fluency and productive vocabulary knowledge during a stay abroad. *Journal of the European Second Language Association*, *5*(1), 101–114. https://doi.org/10.22599/jesla.80

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high stakes outcomes. *Scientific Studies of Reading*, *5*(3), 257–288. https://doi.org/10.1207/S1532799XSSR0503_4

Gough, P., & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10. https://doi.org/10.1177/074193258600700104

Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistic perspective*. Routledge.

Graham, S. (2020). The sciences of reading and writing must become more fully integrated. *Reading Research Quarterly*, *55*(S1), 35–44. https://doi.org/10.1002/rrq.332

Graham, S., & Harris, K. R. (2018). An examination of the design principles underlying a Self-Regulated Strategy Development study. *Journal of Writing Research*, *10*(2), 139–187. https://doi.org/10.17239/jowr-2018.10.02.02

Graham, S., & Hebert, M. (2011). Writing-to-read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, *81*(4), 710–744. https://doi.org/10.17763/haer.81.4.t2k0m13756113566

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., Barkel, A., Kavanaugh, C., & Talukdar, J. (2017). Reading for writing: A meta-analysis of the impact of reading interventions on writing. *Review of Educational Research*, *88*(2), 243–284. https://doi.org/10.3102/ 0034654317746927

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development*, *6*(6), 467–477. https://doi.org/10.1080/01434632.1985.9994221

Grünke, M., & Bracht, J. (2025). Förderung bei Lernstörungen [Support for learning disabilities]. In G. Quenzel, K. Hurrelmann, J. Groß Ophoff & C. Weber (Eds.), *Handbuch Bildungsarmut* [Handbook on educational poverty] (chapter 40–1). Springer VS. https://doi.org/10.1007/978-3-658-44698-7_40-1

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. https://doi.org/10.1037/spq0000199

Haase, A., & Steinbrink, C. (2022). Associations between morphological awareness and literacy skills in German primary school children: The roles of grade level, phonological processing and vocabulary. *Reading and Writing: An Interdisciplinary Journal*, *35*(7), 1675–1709. https://doi.org/10.1007/s11145-021-10247-1

Häfner, I., Flunger, B., Dicke, A.-L., Gaspard, H., Brisson, B. M., Nagengast, B., & Trautwein, U. (2017). Robin Hood effects on motivation in math: Family interest moderates the effects of relevance interventions. *Developmental Psychology*, *53*(8), 1522–1539. https://doi.org/10.1037/dev0000337

Hall, K. J., Fawcett, E. J., Hourihan, K. L., & Fawcett, J. M. (2021). Emotional memories are (usually) harder to forget: A meta-analysis of the item-method directed forgetting literature. *Psychonomic Bulletin & Review*, *28*, 1313–1326. https://doi.org/10.3758/s13423-021-01914-z

Hall, W. (2009). *Dyslexia in the primary classroom*. Learning Matters Ltd.

Harris, K. R., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation* (2nd ed.). Brookline Books.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H. Brookes Publishing Company.

Hasenäcker, J., & Schroeder, S. (2022). Specific predictors of length and frequency effects in German beginning readers: Testing component processes of sublexical and lexical reading in the DRC. *Reading and Writing*, *35*, 1627–1650. https://doi.org/10.1007/s11145-021-10251-5

Hayes, J. R., & Flower, L. S. (1980). Identifying the organisation of writing processes. In L. W. Gregg, & E. R.; Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Lawrence Erlbaum.

Heidari, K. (2024). The contribution of willingness to communicate to L2 learners' depth of vocabulary knowledge: An empirical study. *Journal of Psycholinguistic Research*, *53*, article 35. https://doi.org/10.1007/s10936-024-10062-z

Hjetland, H. N., Hofslundsengen, H., Klem, M., Karlsen, J., Hagen, Å. M., Engevik, L. I., Geva, E., Norbury, C., Monsrud, M. B., & Næss, K. B. (2023). Vocabulary interventions for second language (L2) learners up to six years of age. *The Cochrane Database of Systematic Reviews*, *8*(8), article CD014890. https://doi.org/10.1002/14651858.CD014890.pub2

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*(2), 297–315. https://doi.org/10.1037/a0022128

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological Science*, *23*(6), 572–577. https://doi.org/10.1177/0956797611435921

Institut zur Qualitätsentwicklung im Bildungswesen [IQB]. (2022). *IQB-Bildungstrends 2021: Kompetenzen und schulische Lernbedingungen im zweiten Pandemie-Jahr* [IQB education trends 2021: Competencies and school learning conditions in the second year of the pandemic]. Waxmann.

Isbell, R., Sobol, J., Lindauer, L., & Lowrance, A. (2004). The effects of storytelling and story reading on the oral language complexity and story comprehension of young children. *Early Childhood Education Journal*, *32*(3), 157–163. https://doi.org/10.1023/B:ECEJ.0000048967.94189.a3

Junk-Deppenmeier, A., & Jeuk, S. (Eds.) (2015). *Praxismaterial Förderdiagnostik. Werkzeuge für den Sprachunterricht in der Sekundarstufe I* [Practical material for diagnostic assessment. Tools for language teaching in lower secondary education]. Klett.

Junker, F. B., Sokolowski, H. M., Schlaffke, L., Heba, S., Rüther, N. N., & Schmidt-Wilcke, T. (2024). The angular gyrus serves as an interface between the non-lexical reading network and the semantic system: Evidence from dynamic causal modeling. *Brain Structure and Function*, *229*, 561–575. https://doi.org/10.1007/s00429-023-02624-z

Kang, E. Y., & Shin, M. (2019). The contributions of reading fluency and decoding to reading comprehension for struggling readers in fourth grade. *Reading & Writing Quarterly*, *35*(3), 179–192. https://doi.org/10.1080/10573569.2018.1521758

Karipidis, I. I., Pleisch, G., Brandeis, D., Roth, A., Röthlisberger, M., Schneebeli, M., Walitza, S., & Brem, S. (2018). Simulating reading acquisition: The link between reading outcome and multimodal brain signatures of letter–speech sound learning in prereaders. *Scientific Reports*, *8*, article 7121. https://doi.org/10.1038/s41598-018-24909-8

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–71). Lawrence Erlbaum Associates, Inc.

Kim, Y.-S. G., Yang, D., Reyes, M., & Connor, C. (2021a). Writing instruction improves students' writing skills differentially depending on focal instruction and children: A meta-analysis for primary grade students. *Educational Research Review*, *34*, article 100408. https://doi.org/10.1016/j.edurev.2021.100408

Kim, Y.-S. G., & Zagata, E. (2024). Enhancing reading and writing skills through systematically integrated instruction. *The Reading Teacher*, *77*(6), 787–799. https://doi.org/10.1002/trtr.2307

Kim, Y.-S. G., Quinn, J. M., & Petscher, Y. (2021b). What is text reading fluency and is it a predictor or an outcome of reading comprehension? A longitudinal investigation. *Developmental Psychology*, *57*(5), 718–732. https://doi.org/10.1037/dev0001167

Kittle, J. M., Amendum, S. J., & Budde, C. M. (2024). What does research say about the science of reading for K-5 multilingual learners? A systematic review of systematic reviews. *Educational Psychology Review*, *36*, article 108. https://doi.org/10.1007/s10648-024-09942-6

Klieme, E. (2006). *Zusammenfassung zentraler Ergebnisse der DESI-Studie* [Summary of key findings from the DESI study]. Deutsches Institut für Internationale Pädagogische

Forschung (DIPF). https://www.kmk.org/fileadmin/veroeffentlichungen_beschlu-esse/2006/2006_03_01-DESI-Ausgewaehlte-Ergebnisse.pdf

Koponen, T., Aro, M., Poikkeus, A. M., Niemi, P., Lerkkanen, M. K., Ahonen, T., & Nurmi, J. E. (2018). Comorbid fluency difficulties in reading and math: Longitudinal stability across early grades. *Exceptional Children*, *84*(3), 298–311. https://doi.org/10.1177/0014402918756269

Kraal, A., Koornneef, A. W., Saab, N., & van den Broek, P. W. (2018). Processing of expository and narrative texts by low- and high-comprehending children. *Reading and Writing*, *31*, 2017–2040. https://doi.org/10.1007/s11145-017-9789-2

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

Krieglstein, F., Schneider, S., Beege, M., & Rey, G. D. (2022). How the design and complexity of concept maps influence cognitive learning processes. *Educational Technology Research and Development*, *70*, 99–118. https://doi.org/10.1007/s11423-022-10083-2

Kwakkel, H., Droop, M., Verhoeven, L., & Segers, E. (2024). How lexical quality predicts L2 reading comprehension in early bilingual education. *International Journal of Bilingual Education and Bilingualism*, *27*(7), 993–1007. https://doi.org/10.1080/13670050.2024.2317860

Lange, I., & Gogolin, I. (2010). *Durchgängige Sprachbildung: Eine Handreichung* [Consistent language education: A guide]. Waxmann.

Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316–323). Multilingual Matters.

Law, J., Clegg, J., Rush, R., Roulstone, S., & Peters, T. J. (2019). Association of proximal elements of social disadvantage with children's language development at 2 years: An analysis of data from the Children in Focus (CiF) sample from the ALSPAC birth cohort. *International Journal of Language & Communication Disorders*, *54*(3), 362–376. https://doi.org/10.1111/1460-6984.12442

Ledford, J. R., & Gast, D. L. (Eds.). (2024). *Single case research methodology: Applications in special education and behavioral sciences* (4th ed.). Routledge. https://doi.org/10.4324/9781003294726

Lee, S. (2025). The relationship between receptive and productive knowledge of L2 English collocations. *International Journal of Applied Linguistics*, *35*(1), 109–133. https://doi.org/10.1111/ijal.12605

Lee, Y., Cho, Y.-H., Park, T., & Choi, J. (2023). Situational interest during individual reading and peer reading activities using PALS: Its relationship to students' reading skills and reading motivations. *Educational Studies*, *49*(5), 842–860. https://doi.org/10.1080/03055698.2021.1898341

Lenhard, W., Lenhard, A., & Schneider, W. (2020). *ELFE II: Ein Leseverständnistest für Erst-bis Siebtklässler – Version II* [ELFE II: A reading comprehension test for first to seventh graders - Version II] (4th ed.). Hogrefe.

Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry*, *51*(5), 612–620. https://doi.org/10.1111/j.1469-7610.2009.02185.x

Lesaux, N. K., Rupp, A. A., & Siegel, L. S. (2007). Growth in reading skills of children from diverse linguistic backgrounds: Findings from a 5-year longitudinal study. *Journal of Educational Psychology*, *99*(4), 821–834. https://doi.org/10.1037/0022-0663.99.4.821

Li, S. (2023). Working memory and second language writing: A systematic review. *Studies in Second Language Acquisition*, 45(3), 647–679. https://doi.org/10.1017/S0272263123000189

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

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, *80*(3), 300–335. https://doi.org/10.3102/0034654310377087

Massler, U., Müller, W., Iurgel, I., Haake, S., Gantikow, A., & Hadzilacos, T. (2022). Meaningful, gamified training of reading fluency. *Frontiers in Computer Science*, *4*, article 968137. https://doi.org/10.3389/fcomp.2022.968137

Mathias, B., Sureth, L., Hartwigsen, G., Macedonia, M., Mayer, K. M., & von Kriegstein, K. (2021). Visual sensory cortices causally contribute to auditory word recognition following sensorimotor-enriched vocabulary training. *Cerebral Cortex (New York, N.Y. : 1991)*, *31*(1), 513–528. https://doi.org/10.1093/cercor/bhaa240

Mayer, A. (2020). *Test zur Erfassung der phonologischen Bewusstheit und der Benennungsgeschwindigkeit (TEPHOBE): Manual* [Test for the assessment of phonological awareness and rapid naming (TEPHOBE): Manual] (4th ed.). Ernst Reinhardt Verlag. http://www.reinhardt-verlag.de/de/titel/54688/

McKeown, M. G. (2019). Effective vocabulary instruction fosters knowing words, using words, and understanding how words work. *Language, Speech, and Hearing Services in Schools*, *50*(4), 466–476. https://doi.org/10.1044/2019_LSHSS-VOIA-18-0126

McLeskey, J., Barringer, M.-D., Billingsley, B., Brownell, M., Jackson, D., Kennedy, M., & Ziegler, D. (2017). *High-leverage practices in special education*. Council for Exceptional Children & CEEDAR Center.

Moeyaert, M., Klingbeil, D. A., Rodabaugh, E., & 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 and Special Education*, *42*(2), 94–106. https://doi.org/10.1177/0741932519855079

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524759

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139858656

National Reading Panel (2000). *Teaching children to read.* National Institute of Child Health and Human Development.

Neuman, S. B., Kaefer, T., & Pinkham, A. (2022). Incidental word learning through multiple media: A case for synergy. *Topics in Language Disorders*, *42*(4), 290–303. https://doi.org/10.1097/TLD.0000000000000293

Nouwens, S., Groen, M. A., Kleemans, T. & Verhoeven, L. (2021). How executive functions contribute to reading comprehension. *British Journal of Educational Psychology*, *91*(1), 169–192. https://doi.org/10.1111/bjep.12355

Okkinga, M., van Steensel, R., van Gelderen, A. J. S., van Schooten, E., Sleegers, P. J. C., & Arends, L. R. (2018). Effectiveness of reading-strategy interventions in whole classrooms: A meta-analysis. *Educational Psychology Review*, *30*(4), 1215–1239. https://doi.org/10.1007/s10648-018-9445-7

Organisation for Economic Co-operation and Development [OECD] (Ed.). (2023). *PISA 2022 Ergebnisse: Lernstände und Bildungsgerechtigkeit* [PISA 2022 Results: Learning Levels and Educational Equity]. Waxmann. https://doi.org/10.3278/6004956w

Organisation for Economic Co-operation and Development [OECD] (Ed.). (2019). *PISA 2018 Ergebnisse: Lernstände und Bildungsgerechtigkeit* [PISA 2018 Results: Learning Levels and Educational Equity]. Waxmann. https://doi.org/10.31244/9783830991007

Organisation for Economic Co-operation and Development [OECD]. (2023). *PISA 2022 Results — Factsheets: Germany* (pp. 5–6). OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/11/pisa-2022-results-volume-i-and-ii-country-notes_2fca04b9/germany_bf0846ee/1a2cf137-en.pdf

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*(3), 255–287. https://doi.org/10.1037/h0084295

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*(2), 117–175. https://doi.org/10.1207/s1532690xci0102_1

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

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

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011b). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, *42*(2), 284–299. https://doi.org/10.1016/j.beth.2010.08.006

Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, *8*(3), 317–344. https://doi.org/10.1016/0361-476X(83)90019-X

Perfetti, C., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro & P. Reitsma (Ed.), *Precursors of Functional Literacy* (pp. 189–213). John Benjamins Publishing Company. https://doi.org/10.1075/swll.11.14per

Petermann, F. (2018). *Sprachstandserhebungstest für Kinder im Alter zwischen 5 und 10 Jahren* [Language assessment test for children between the ages of 5 and 10] (3rd ed.). Hogrefe.

Phillips Galloway, E., & Uccelli, P. (2019). Examining developmental relations between core academic language skills and reading comprehension for English learners and their peers. *Journal of Educational Psychology*, *111*(1), 15–31. https://doi.org/10.1037/edu0000276

Phillips Galloway, E., McClain, J. B., & Uccelli, P. (2020). Broadening the lens on the science of reading: A multifaceted perspective on the role of academic language in text understanding. *Reading Research Quarterly*, *55*(S1), 331–345. https://doi.org/10.1002/rrq.359

Rehfeld, D. M., Kirkpatrick, M., O'Guinn, N., & Renbarger, R. (2022). A meta-analysis of phonemic awareness instruction provided to children suspected of having a reading disability. *Language, Speech, and Hearing Services in Schools*, *53*(4), 1177–1201. https://doi.org/10.1044/2022_LSHSS-21-00160

Rohm, T., Carstensen, C. H., Fischer, L., & Gnambs, T. (2021). The achievement gap in reading competence: The effect of measurement non-invariance across school types. *Large-scale Assessments in Education*, *9*, article 23. https://doi.org/10.1186/s40536-021-00116-2

Röthlisberger, M., Zangger, C., & Juska-Bacher, B. (2023). The role of vocabulary components in second language learners' early reading comprehension. *Journal of Research in Reading*, *46*(1), 1–21. https://doi.org/10.1111/1467-9817.12411

Sachverständigenrat deutscher Stiftungen für Integration und Migration (SVR). (2013). *Hürdenlauf zur Kita: Warum viele Kinder mit Migrationshintergrund von frühkindlicher Bildung ausgeschlossen sind* [Hurdles to kindergarten: Why many children with a migration background are excluded from early childhood education]. SVR GmbH. https://www.svr-migration.de/publikationen/huerdenlauf-zur-kita/

Sáenz, L. M., Fuchs, L. S., & Fuchs, D. (2005). Peer-assisted learning strategies for English language learners with learning disabilities. *Exceptional Children*, *71*(3), 231–247. https://doi.org/10.1177/001440290507100302

Sanders, S. (2020). Using the self-regulated strategy development framework to teach reading comprehension strategies to elementary students with disabilities. *Education and Treatment of Children*, *43*(1), 57–70. https://doi.org/10.1007/s43494-020-00009-z

Sanders, S., Losinski, M., Parks Ennis, R., White, W., Teagarden, J., & Lane, J. (2019). A meta-analysis of self-regulated strategy development reading interventions to improve the reading comprehension of students with disabilities. *Reading & Writing Quarterly*, *35*(4), 339–353. https://doi.org/10.1080/10573569.2018.1545616

Sanetti, L. M. H., Cook, B. G., & Cook, L. (2021). Treatment fidelity: What it is and why it matters. *Learning Disabilities Research & Practice*, *36*(1), 5–11. https://doi.org/10.1111/ldrp.12238

Sappok, C., Linnemann, M., & Stephany, S. (2020). Leseflüssigkeit – Prosodie – Leseverstehen. Eine Longitudinalstudie zur Entwicklung der Leseflüssigkeit von Jahrgangstufe 3 bis 7 [Reading fluency – prosody – reading comprehension. A longitudinal study on the development of reading fluency from grades 3 to 7]. In I. Rautenberg (Ed.), *Evidenzbasierte Forschung zum Schriftspracherwerb* [Evidence-based research on written language acquisition] (pp. 175–209). Schneider Verlag Hohengehren.

Schoonen, R. (2019) Are reading and writing building on the same skills? The relationship between reading and writing in L1 and EFL. *Reading & Writing*, *32*, 511–535. https://doi.org/10.1007/s11145-018-9874-1

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology*, *60*, article 101832. https://doi.org/10.1016/j.cedpsych.2019.101832

Sehlström, P., Waldmann, C., & Levlin, M. (2023) Self-efficacy for writing and written text quality of upper secondary students with and without reading difficulties. *Frontiers in Psychology*, *14*, article 1231817. https://doi.org/10.3389/fpsyg.2023.1231817

Seifert, S., Paleczek, L., Schwab, S., & Gasteiger-Klicpera, B. (2017). *Grazer Wortschatztest* [Graz vocabulary test]. Hogrefe.

Sencibaugh, J. M. (2007). Meta-analysis of reading comprehension interventions for students with learning disabilities: Strategies and implications. *Reading Improvement*, *44*(1), 6–22.

Share D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151–226. https://doi.org/10.1016/0010-0277(94)00645-2

Shin, J. (2020). A meta-analysis of the relationship between working memory and second language reading comprehension: Does task type matter? *Applied Psycholinguistics*, *41*(4), 873–900. https://doi.org/10.1017/S0142716420000272

Skar, G. B., Lei, P. W., Graham, S., Aasen, A. J., Byberg Johansen, M., & Holten Kvistad, A. (2022). Handwriting fluency and the quality of primary grade students' writing. *Reading and Writing*, *35*(2), 509–538. https://doi.org/10.1007/s11145-021-10185-y

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension.* RAND Corporation. https://www.rand.org/pubs/monograph_reports/MR1465.html

Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology*, *93*(4), 735–749. https://doi.org/10.1037/0022-0663.93.4.735

Spörer, N., Seuring, V., Schünemann, N., & Brunstein, J. C. (2009). Fostering the reading comprehension of secondary school students through peer-assisted learning: Effects on strategy knowledge, strategy use, and task performance. *Contemporary Educational Psychology*, *34*(4), 289–297. https://doi.org/10.1016/j.cedpsych.2009.06.004

Statistisches Bundesamt (Destatis). (2024a). *Bevölkerung in Privathaushalten 2024 nach Migrationshintergrund (Mikrozensus-Auswertung)* [Population in private households in 2024 by migration background (microcensus evaluation)]. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Tabellen/migrationshintergrund-alter.html

Statistisches Bundesamt (Destatis) (2024b). *Sozialbericht 2024. Ein Datenreport für Deutschland* [Social Report 2024: A data report for Germany]. Bundeszentrale für politische Bildung (bpb). https://www.bpb.de/kurz-knapp/zahlen-und-fakten/sozialbericht-2024/

Stevens, E. A., Park, S., & Vaughn, S. (2018). A review of summarizing and main idea interventions for struggling readers in grades 3 through 12: 1978–2016. *Remedial and Special Education*, *40*(3), 131–149. https://doi.org/10.1177/0741932517749940

Sun, Y., Wang, J., Dong, Y., Zheng, H., Yang, J., Zhao, Y., & Dong, W. (2021). The relationship between reading strategy and reading comprehension: A meta-analysis. *Frontiers in Psychology*, *12*, article 635289. https://doi.org/10.3389/fpsyg.2021.635289

Taha, J., Carioti, D., Stucchi, N., Chailleux, M., Granocchio, E., Sarti, D., De Salvatore, M., & Guasti, M. T. (2022). Identifying the risk of dyslexia in bilingual children: The potential of language-dependent and language-independent tasks. *Frontiers in Psychology*, *13*, article 935935. https://doi.org/10.3389/fpsyg.2022.935935

Tahmouresi, S., & Papi, M. (2021). Future selves, enjoyment and anxiety as predictors of L2 writing achievement. *Journal of Second Language Writing*, *53*, article 100837. https://doi.org/10.1016/j.jslw.2021.100837

Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2018). I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences*, *61*, 136–150. https://doi.org/10.1016/j.lindif.2017.11.015

Tang, S., Irby, B. J., Tong, F., & Lara-Alecio, R. (2021). The effects of cooperative, collaborative, and peer-tutoring strategies on English learners' reading and speaking proficiencies in an English-medium context: A research synthesis. *SAGE Open*, *11*(4), article 21582440211060823. https://doi.org/10.1177/21582440211060823

Tenenbaum, H. R., Winstone, N. E., Leman, P. J., & Avery, R. E. (2020). How effective is peer interaction in facilitating learning? A meta-analysis. *Journal of Educational Psychology*, *112*(7), 1303–1319. https://doi.org/10.1037/edu0000436

Thordardottir, E., & Plez, L. (2024). The effect of age of first exposure on vocabulary, mean length of utterance, morphosyntactic accuracy, and semantic and sentence-level patterns in the first 2 years of French second-language learning by preschool- to adolescent-age Mandarin speakers. *Journal of Speech, Language, and Hearing Research*, *67*(9), 3040–3063. https://doi.org/10.1044/2024_JSLHR-23-00568

Thurston, A., Cockerill, M., & Chiang, T.-H. (2021). Assessing the differential effects of peer tutoring for tutors and tutees. *Education Sciences*, *11*(3), 97. https://doi.org/10.3390/educsci11030097

Toste, J. R., Didion, L., Peng, P., Filderman, M. J., & McClelland, A. M. (2020). A meta-analytic review of the relations between motivation and reading achievement for K–12 students. *Review of Educational Research*, *90*(3), 420–456. https://doi.org/10.3102/0034654320919352

Troia, G. A., & Graham, S. (2002). The effectiveness of a highly explicit, teacher-directed strategy instruction routine: Changing the writing performance of students with learning disabilities. *Journal of Learning Disabilities*, *35*(4), 290–305. https://doi.org/10.1177/00222194020350040101

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlation studies. *Language Learning*, *69*(3), 559–599. https://doi.org/10.1111/lang.12343

Valcárcel Jiménez, M., Yumus, M., Schiele, T., Mues, A., & Niklas, F. (2024). Preschool emergent literacy skills as predictors of reading and spelling in Grade 2 and the role of migration background in Germany. *Journal of Experimental Child Psychology*, *244*, article 105927. https://doi.org/10.1016/j.jecp.2024.105927

van de Pol, J., Mercer, N., & Volman, M. (2019). Scaffolding student understanding in small-group work: Students' uptake of teacher support in subsequent small-group interaction. *Journal of the Learning Sciences*, *28*(2), 206–239. https://doi.org/10.1080/10508406.2018.1522258

Verhoeven, L., & Perfetti, C. (2021). Universals in learning to read across languages and writing systems. *Scientific Studies of Reading*, *26*(2), 150–164. https://doi.org/10.1080/10888438.2021.1938575

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 of German schoolchildren. *Journal of School Psychology*, *66*, 25–40. https://doi.org/10.1016/j.jsp.2017.11.003

Vygotsky, L. S. (1978). Interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind and society: The development of higher psychological processes* (pp. 79–91). Harvard University Press.

Walter, J. (2009). *LDL Lernfortschrittsdiagnostik Lesen – Ein curriculumbasiertes Verfahren* [LDL Reading Progress Diagnostics – A Curriculum-Based Method]. Hogrefe.

Walter, J. (2013). *VSL: Verlaufsdiagnostik sinnerfassenden Lesens* [VSL: Progress diagnostics for reading comprehension]. Hogrefe.

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.

Weiß, R. H. (2007). *WS/ZF-R. Wortschatztest und Zahlenfolgentest – Revision. Ergänzungstest zum CFT 20-R* [WS/ZF-R. Vocabulary test and number sequence test — Revision. Supplementary test to the CFT 20-R]. Hogrefe.

Wilbert, J., & Lueke, T. (2021). *Package 'scan'*. https://cran.r-project.org/web/packages/scan/scan.pdf

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *Journal of Speech, Language, and Hearing Research*, *50*(4), 1093–1109. https://doi.org/10.1044/1092-4388(2007/076)

Wright, A. (2013). *RBT: Creating stories with children*. Oxford University Press.

Wright, T. S., & Cervetti, G. N. (2017). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly*, *52*(2), 203–226. https://doi.org/10.1002/rrq.163

Wu, H. (2022). A brief review of the effect of reading strategy instruction on L2 reading comprehension: A discussion of types, amount, L2 proficiency, and age. *Journal of Language Teaching*, *2*(3), 1–5. https://doi.org/10.54475/jlt.2022.003

Yurtsever, A., Anderson, J. A. E., & Grundy, J. G. (2023). Bilingual children outperform monolingual children on executive function tasks far more often than chance: An updated quantitative analysis. *Developmental Review*, *69*, article 101084. https://doi.org/10.1016/j.dr.2023.101084

Zhang, J., & Wang, Q. (2023). Reading and writing difficulties in bilingual learners. *Annals of Dyslexia*, *73*, 1–5. https://doi.org/10.1007/s11881-023-00282-8

Zhang, S., & Zhang, X. (2022). The relationship between vocabulary knowledge and L2 reading/listening comprehension: A meta-analysis. *Language Teaching Research*, *26*(4), 696–725. https://doi.org/10.1177/1362168820913998

Appendix A Article 1 *(peer-reviewed)*

Disclaimer for Dissertation Version:

Abstract

The acquisition of reading skills in a person's native language is a key life component to functioning successfully within their society. Unfortunately, many German students have significant reading difficulties in their native language, German. Therefore, the purpose of this study is to translate into German and implement a promising U.S. reading programme that focuses on fluency and comprehension: Peer-Assisted Learning Strategies (PALS Reading 2–6). PALS has shown positive effects for English elementary and secondary students with or at risk for reading disabilities. Therefore, this project conducted a pilot Single Subject multiple baseline design study with eight 7th grade students with or at risk for reading disabilities, to explore the feasibility of implementing PALS in German. Results showed improvement in German reading comprehension and fluency for these students after five sessions along with stability at follow-up conducted after six weeks. Limitations of the study and areas for further research are discussed, as well as implications for practitioners.

Keywords: PALS; Reading Comprehension; Reading Fluency; Single Case; Peer Tutoring; Students with or At-Risk for Reading Disabilities (SRWD)

Introduction

Reading proficiency is essential to many areas of life (Grigoryan 2020). Deficits in reading are linked to academic, behavioural, and socioeconomic challenges (Bennett et al. 2003). Consequently, it is a core mandate of society that every child acquire the necessary reading skills (McElvany et al. 2023). Despite its importance, there is a noticeable decline in basic reading

skills both nationally and internationally. Studies report that 25% of students are classified as having or being at-risk for reading disabilities (SWRD). These students are reading at or below the Basic level, which refers to a standard of reading proficiency where students can perform fundamental reading tasks but struggle with more advanced skills (Garwood, Brunsting, and Fox 2014; McElvany et al. 2023). SWRD exhibit significant difficulties in acquiring age-appropriate reading skills and experience difficulties in various reading domains, including fluency and comprehension. The 2022 Programme for International Student Assessment (PISA) study in Germany yielded analogous results for secondary school, with 25% of fifteen-year-olds SWRD. Over the past 20 years, average reading proficiency in Germany has deteriorated, with an increasing proportion of low-achieving students (Organization for Economic Cooperation and Development (OECD) 2023).

Reading development is influenced by language skills, background factors, and reading practice, evolving with experience and age. While early reading difficulties are primarily characterised by problems with word reading (Snowling, Hulme, and Nation 2020), older students may face challenges in both foundational and advanced skills, such as text comprehension (Hjetland et al. 2020). As students transition from 'learning to read' to 'reading to learn', typically in upper elementary and beyond, they are expected to comprehend increasingly complex texts (Wanzek et al. 2010). Although this transition begins at the end of primary school, the demands on students' reading skills increase and students in secondary school receive less explicit support in reading. This shift highlights the need for tailored support for secondary students, especially for SWRD, who may struggle with both decoding and comprehending challenging material (Snowling, Hulme, and Nation 2020).

Effective reading interventions for secondary SWRD

Research has identified several components of effective reading programmes (Calhoon, Sandow, and Hunter 2010; Wanzek et al. 2010). For secondary SWRD, meta-analyses have shown that targeted reading interventions significantly improve reading outcomes (e.g. Filderman et al. 2022; Roberts et al. 2020). Interventions focusing on reading fluency, comprehension, and mixed approaches have proven to be highly effective and are often supported by strategy instruction (e.g. Filderman et al. 2022; Suggate 2010, 2016). For example, Donegan and Wanzek (2021) found that multicomponent interventions were the only type of intervention to produce significant effects on both comprehension and foundational outcomes.

To be most effective, interventions for SWRD should be systematic, explicit, and tailored to individual needs (e.g. Bakken et al. 2021; Graham, Silva, and Restrepo 2023; Hall et al. 2022; Kim et al. 2020; Scammacca et al. 2016). Small group instruction, which provides

more personalised support, has been shown to be particularly beneficial (Donegan and Wanzek 2021; Kim et al. 2020). Furthermore, long-term, more intense and high-dosage interventions yielded stronger effects (Hall et al. 2022; Kim et al. 2020). Moreover, motivation-focused interventions (Van Der Sande et al. 2023) as well as peertutoring procedures (students alternate as tutors and tutees) (Bowman-Perrott et al. 2013; Moeyaert et al. 2021; see also Völlinger, Supanc, and Brunstein 2018) seem to be beneficial. Despite the clear benefits of multicomponent programmes, there has been little research for secondary SWRD (Wanzek et al. 2010; Donegan and Wanzek 2021).

Peer-assisted-learning strategies

One such multicomponent programme is the Peer-Assisted Learning Strategies (PALS) reading programme, developed by Fuchs et al. (1997). PALS enhances reading abilities through structured activities and peer-assisted learning sessions, fostering positive peer interactions and personalised instruction by pairing students of varying skills (Fuchs, Fuchs, and Abramson 2020). The PALS programme incorporates four essential strategic activities: Partner Reading (students taking turns reading aloud and coaching each other for 5 minutes each, using correction procedures for mistakes), Retell (one student summarising the text they read together for 2 minutes, with the other student correcting any inaccuracies), Paragraph Shrinking (one student reading a paragraph and summarising the main idea while the other student provides prompts and corrections), and Prediction Relay (one student making predictions about the text and verifying them after reading, with roles switching every 5 minutes).

Research found that PALS is an appropriate, culturally adaptable multicomponent programme that enhances reading skills among students with diverse needs. Lee et al. (2023) highlighted its effectiveness for students with lower reading abilities and motivation. Fuchs et al. (1997, 2021) demonstrated PALS' success in boosting various literacy skills. Additionally, Sáenz et al. (2005) and Harsul (2022) noted its effectiveness across different languages and diverse student populations. Calhoon and Fuchs (2003, 2005, 2010) reported significant gains in comprehension for middle school SWRD reading.

In German speaking countries, however, little is known about PALS. Vardy et al. (2022) emphasised the need for adaptations to different cultural contexts and languages. In secondary education, only Spörer et al. (2009) employed PALS to enhance reading abilities in 7th grade students. In this study seventy-four students were divided into three groups: (a) German training with German materials, (b) English training with English materials, and (c) a control group with no training. Evaluations using reading comprehension tests showed that students in both training conditions performed better in the post-test and follow-up than those in the control group.

Notably, students in the German training group outperformed those in the English training group. Additionally, improvements were observed in fluent reading and text summarisation during the training. However, these results do not specifically relate to SWRD.

Given that PALS Reading 2–6 is appropriate for the beginning of secondary school in Germany because secondary education starts after the fourth grade, and particularly SWRD in the first years of secondary education have learning levels comparable to younger students, it is necessary to adapt and evaluate this programme in the German context. This led us to the research question:

- (a) Does an adapted German-language version of PALS improve reading fluency skills for secondary SWRD?
- (b) Does an adapted German-language version of PALS improve reading comprehension skills for secondary SWRD?
- (c) How was the adapted German-language PALS intervention evaluated in terms of social validity by secondary SWRD?

Methods

Participants and setting

The study was conducted at an urban secondary school in North Rhine-Westphalia, Germany. Informed consent forms were sent to all legal guardians before the study began. The class teacher of a 7th grade class agreed to the project because he has several SWRD. To select participants for the study, the reading skills of all students (N = 23) in this 7th grade class were assessed, and those needing support in reading comprehension (VSL and ELFE percentile rank < 25) were selected, resulting in twelve participants (N = 12). As the baseline session should not include less than three sessions (Gast, Lloyd, and Ledford 2018), and four missing values in the intervention also led to exclusion, the number of participants was reduced to eight by COVID-19 due to absenteeism (see Table 1).

Screenings

Students were tested either individually or in groups before the start of the study (depending on the screening procedure). Extra rooms were used for this purpose so that the screenings could be carried out quietly. The twelve participants were divided into three small groups with 3 baselines of different lengths (see design chapter), with two students working together in pairs in each small group.

ITRF-G. Reading and behavioural difficulties can negatively impact each other (Garwood, Brunsting, and Fox 2014). To assess problem behaviour, the German Integrated Teacher Report Form-German (ITRF-G, Volpe et al. 2018) was used to obtain additional information

about the effectiveness of the intervention for individual students. It includes academic productivity problems/disorganisation and oppositional/disruptive behaviour. The ITRF-G has high external validity, acceptable test-retest reliability and high internal consistency (α = .91), as reported by Volpe et al. (2018).

ELFE-II. To assess reading comprehension at word, sentence, and text levels the ELFE II (1–7th grade) was used as a screening tool. This test was administered as a paper-and-pencil test in a group setting for 20 to 30 minutes. Lenhard, Lenhard, and Schneider (2020) report an odd-even split-half reliability for the total score on the paper form (N = 1520) of rtt=.096, a validity as measured by the correlation with another standardised reading test (SLS 2–9) of rct = .77, and agreement with the teacher's assessment of reading performance of rct = .70.

Table 1. Demographic Characteristics of the Participants.

Name	Gender	Age	Migra-tion	Lingua Franca German	SEN	VSL	LDL	ELFE II				ITRF-G
								W	S	T	O	
B03	female	14	Yes	No	LD	0	34	36	34	37	8.08*	0
B05	male	13	Yes	No		0	38	36	38	46	15.87*	19*
B07	male	14	No	Yes		2	27	43	46	41	24.2	4
B08	male	13	Yes	No	SP	0	7	42	37	41	15.87*	22*
B09	female	14	No	Yes		10	73	39	47	44	24.2	13*
B10	female	13	Yes	No		N.A.	N.A.	37	38	39	11.51*	16*
B11	female	14	Yes	Yes		13	1	39	34	27	4.46*	5
B12	female	13	Yes	No	LD	0	34	42	42	41	21.19	32*

Note. Diagnosed SEN; Learning Disabilities (LD); SEN in speech, language and communication (SP); VSL test reading comprehension, Percentile ranks; Integrated Teacher Report Form German (ITRF-G), *(cut-off ≥13) conspicuous in behaviour; ELFE II test reading, T-values at Word-level (W), Sentence-level (S), and Text-level (T), Percentile rank for overall results (O), *below average reading comprehension; LDL test reading fluency, Percentile ranks; not available (N.A.)

Experimental design

To avoid maturation or intermediate events as alternate explanations, a multiple-baseline across participants with an AB plan was applied (Lane, Ledford, and Gast 2017). Phase A served as the baseline (data collection without reading intervention as control), whereas Phase B comprised the actual reading intervention. Kazdin (2011) recommends using at least three baseline data points. Due to organisational and illness-related problems, some sessions could not be performed. Nevertheless, the study used four baseline sessions for group one, five for group two and six for group three. The intervention phase varied in duration, ranging from five sessions

for group three to seven sessions for group one. Throughout the course of the study, baseline and intervention sessions took place three times per week, with each session lasting 45 minutes.

Dependent variables and data collection

During the study, two dependent variables were tried to manipulate through the reading intervention: 1) reading comprehension and 2) reading fluency. Data were collected by the assessors individually for each student in a separate room.

The reading comprehension progress monitoring (VSL)

The VSL (Walter 2010) is a longitudinal assessment for evaluating reading proficiency. The task is to read a text in which every seventh word contains a bracket with three choice words, i.e. two distractors and the matching word. In each case, the appropriate word must be circled. Because of the large number of parallel forms and the short test time, the VSL can be administered at up to 20 points in time either as a paper-and-pencil or as a PC version. Walter (2010) report an internal consistency of .93. and a construct reliability of H = .94, and a parallel test reliability of rtt.77 to rtt.86. The VSL was used as a screening and to record the dependent variables (reading comprehension) in progress monitoring.

The reading fluency progress monitoring (LDL)

The LDL is an assessment tool for reading fluency and administered as a single test, taking about two minutes to complete. There are 28 reading texts that students must read aloud for one minute at each time point. According to Walter (2010), the parallel test reliability for secondary school (grades 7–9) is rtt=.80 (N = 435) and the criterion-oriented validity was determined on the basis of other reading screenings like the ELFE 1–6: r = .84 and the SLS 1–4: r = .94.

Treatment procedures

Baseline (A phase)

To increase internal validity and counteract reactivity (i.e. the influence that an observer has on the behaviour under observation), the baseline phase entailed activities for the same duration as the reading intervention. The content encompassed different activities (i.e. maths tasks) that did not already foster reading. The student groups remained consistent throughout both the baseline and intervention phases. At the end of each baseline session, all participants were assessed independently with respect to the dependent variables. Due to time constraints, the gradual introduction of the structure of the adapted PALS programme began in the last baseline session. In preparation to learn the reading techniques in this programme, the students learned about the purpose of PALS, the adapted PALS roles: 'coach' and 'athlete', the partner tasks, the teams and team points, the PALS materials, and the four PALS rules: (1) talk only to your partner and

only about PALS; (2) keep your voice low; (3) cooperate with your partner; (4) try your best (Fuchs et al. 2008a).

PALS intervention (B phase)

Material. The PALS programme by Fuchs et al. (1997) was translated into German. For the intervention phase, there were three books to choose from. They were easy to read with themes adapted for youth. Each student received a PALS binder containing the following (1) title page for names and team names (team = two students as a pair); (2) table to record reading scores and the corresponding session; (3) another chart visible to all students for team names and points; (4) a bookmark for students to use with the four PALS rules written on the back; (5) a question card for the partner reading and retelling process; (6) a correction card for dealing with reading errors and paragraph shortening support; (7) information on the predictive reading procedure and an overview of the four reading errors; and (8) certificates awarded upon successful completion as a PALS professional.

PALS procedure. The intervention phase involved the sequential introduction of the four PALS reading strategies across multiple sessions. Initially, each session commenced with a review of the content covered previously, followed by a gradual introduction of reading strategies (partner reading, retell, paragraph shrinking, and prediction relay) over the first five sessions. From the sixth session onwards, the complete PALS programme was implemented. The interventionists prepared the materials, guided students through the procedure, observed students, provided feedback, and distributed points for PALS compliance.

(1) In the first session Partner Reading was introduced, incorporating the role-switching activity between 'athlete' (reading aloud while answering questions), and 'coach' (following along and reading silently with their 'athlete', while helping their partner with mistakes or difficult words), emphasising error corrections (wrong word or word ending, omitting or adding a word or word ending). The subsequent sessions focused on (2) error corrections techniques and retelling strategies (3 & 4) paragraph shrinking, and (5) prediction relay.

At the end of each session, the student's teams were scored. Students could give themselves one point each for following the rules. Additionally, the interventionists could give an additional point to a team. The points were collected in the form of marbles and also recorded on a visible chart. Teams remained unchanged throughout the intervention period. Sessions ended with data collection. Upon completion of all PALS strategies, students received a certificate as an official PALS Reading Professional. All the students went through all the PALS steps. Unfortunately, there were not so many sessions where all students could go through the

entire PALS programme because students were absent from school from time to time (mainly due to COVID).

Differences to the English PALS program. Contrasting the U.S. PALS, this adaptation does not include additional texts for practice. Instead, students are encouraged to read their own books. Additionally, the introductory sessions have been shortened from 12 to 6, allowing students to complete entire PALS sessions more quickly. The reduction in the length of the original programme facilitates transfer and practice, as lengthy programmes would present educators with additional challenges due to disruption of vacation periods. Certain phrases were cancelled due to the interventionists feeling inauthentic with this type of language (e.g. 'Great! Now let's get excited about our next PALS activity' (Fuchs et al. 2008b, 62). The roles of coach and reader were translated as coach and athlete in terms of sports. In the U.S. Pals, students remain First and Second Readers throughout the 4-week training period because the pairs consist of a higher reader and a lower reader (Fuchs et al. 2008a). However, as this study only took place with SWRD, the roles in the adapted version changed at the beginning. Marbles were used to help visualise the points that the students could receive.

Interventionists. Three Master's students of special education acted as interventionists (conducting baseline procedure and intervention) to avoid bias. All interventionists received intensive training in two 2 h-meetings. The interventionists rotated from session to session whereby always two of the interventionists carried out the baseline and intervention sessions together.

Assessors. Two Master's students of special education were responsible only for data collection. They received a 1.5 h-training session on how to assess data properly. Both assessors were always present during the data collection. In addition, they were responsible for collecting treatment fidelity and social validity data.

Treatment fidelity

To increase internal validity and ensure that the intervention was delivered as planned, treatment fidelity was recorded (Sanetti, Cook, and Cook 2021) using a researcher-created checklist, completed by the assessors after each intervention session. The checklist included 19 items (see Online Supplementary Material), each of which could be answered yes or no and were assigned to the following five domains: (1) Environmental Conditions (2) Planning (3) Materials (4) Procedure (5) Diagnostic and Feedback. In addition, comments could be made on specifics within the context of the intervention.

Social validity

Social validity refers to the acceptance and approval of intervention goals, procedures, and outcomes by those receiving and providing services, and other consumers (Luiselli 2021). In the

present study, the Usage Rating Profile – Intervention by Briesch et al. (2013) was used and a questionnaire was created to assess the acceptance, understanding, and feasibility of the intervention among students and teachers. The student questionnaire comprised 16 items formulated as statements (e.g. I liked coming to the intervention.) and could be checked on a five-point Likert scale (0 = 'not true at all' to 4 = 'completely true'). There was space for additional student comments. The class teacher was given a different questionnaire with eight items on the perceived impact of the support on students' behaviour.

Data analysis

The 'scan' package by Wilbert and Lueke (2021) in 'R' was used to analyse data. First, graphs for each dependent variable were created with all participants being shown in one graph to determine trend, level and slope. Descriptive data was calculated. In addition, overlap indices were utilised as follows: the Non-Overlap of All Pairs (NAP; Parker, Vannest, and Davis 2011) and the Tau-U (A vs. B + TrendB – TrendA; Parker et al. 2011) were used. Tau-U was additionally added due to possible trends in each phase to prevent over-estimation of results. Further, the between-case standardised mean difference (BC-SMD; Chen et al. 2023) was calculated using the online calculator. As an estimation method, the restricted maximum likelihood (REML) was used. A fixed and random effect was specified for the baseline level and a fixed and random effect for the intervention level. The BCSMD represents the mean differences between conditions relative to the standard deviation (Chen et al. 2023). It is similar to Hedges'g.

Results

Visual analysis indicated a functional relation between PALS and both dependent variables (see Figures 1, 2). Further, the BC-SMD was 1.21 (CI95 [0.53, 1.89]) for fluency and 2.07 (CI95 [1.05, 3.08]) for comprehension, which is a 1.21 or 2.07 increase in standardised units from the baseline to the intervention phase.

Reading fluency

Overall, the PALS intervention demonstrated improvements in reading fluency for most students, B09, B07, B12, B05, and B10 all showed substantial increases from baseline to intervention, with gains that were either maintained or further improved in the follow-up phase, indicating strong and lasting effects. B11 also showed notable improvement, although there was a slight decrease in follow-up, reflecting moderate to strong effects. B08 presented more moderate improvement with smaller gains compared to other participants. B03 displayed moderate improvements, but these changes were not significant, showing only a weak correlation between baseline and intervention. Overall, seven out of eight students showed significant improvements (see Table 2).

Table 2. Descriptive data in A, B and C phase for reading fluency (LDL).

	N_A	N_B	N_C	M_A (SD)	M_B (SD)	M_C (SD)	MBD (AB)	Max A	Max B	Max C	NAP	Tau-U AB	Hedges' g
B09	4	7	3	30.50 (7.05)	57.71 (6.16)	55.67 (7.77)	27.21	38	66	62	100**	0.40	3.85
B11	4	6	3	20.75 (9.50)	55.50 (11.50)	52.67 (12.86)	34.75	31	73	62	100**	0.38	2.91
B07	4	6	2	27.75 (7.24)	58.00 (26.70)	82.50 (19.09)	30.25	38	93	96	92*	0.67**	1.27
B08	5	6	3	9.40 (3.51)	16.00 (7.65)	27.00 (9.64)	6.6	14	27	34	77	0.46	0.97
B12	4	6	2	27.25 (9.61)	68.50 (25.88)	76.00 (5.66)	41.25	34	93	80	98**	0.69**	1.75
B03	6	5	2	44.67 (8.83)	57.20 (11.10)	62.50 (23.34)	12.53	54	66	79	83*	0.33	1.15
B05	5	5	2	14.20 (7.47)	37.60 (12.10)	37.00 (1.41)	23,4	21	58	38	100**	0.45	2.10
B10	6	5	3	52.00 (15.30)	70.80 (3.19)	85.00 (5.29)	18.8	60	73	89	100**	0.57*	1.48

Note. Measurements exclusive missing values (N); Baseline (A); Intervention (B); Follow-Up (C); Mean (M); Standard Deviation (SD); Mean Baseline Difference (MBD); Maximum (Max); *<.05; **<.01

Figure 1. Visual Analysis in A, B and C phase for reading fluency.
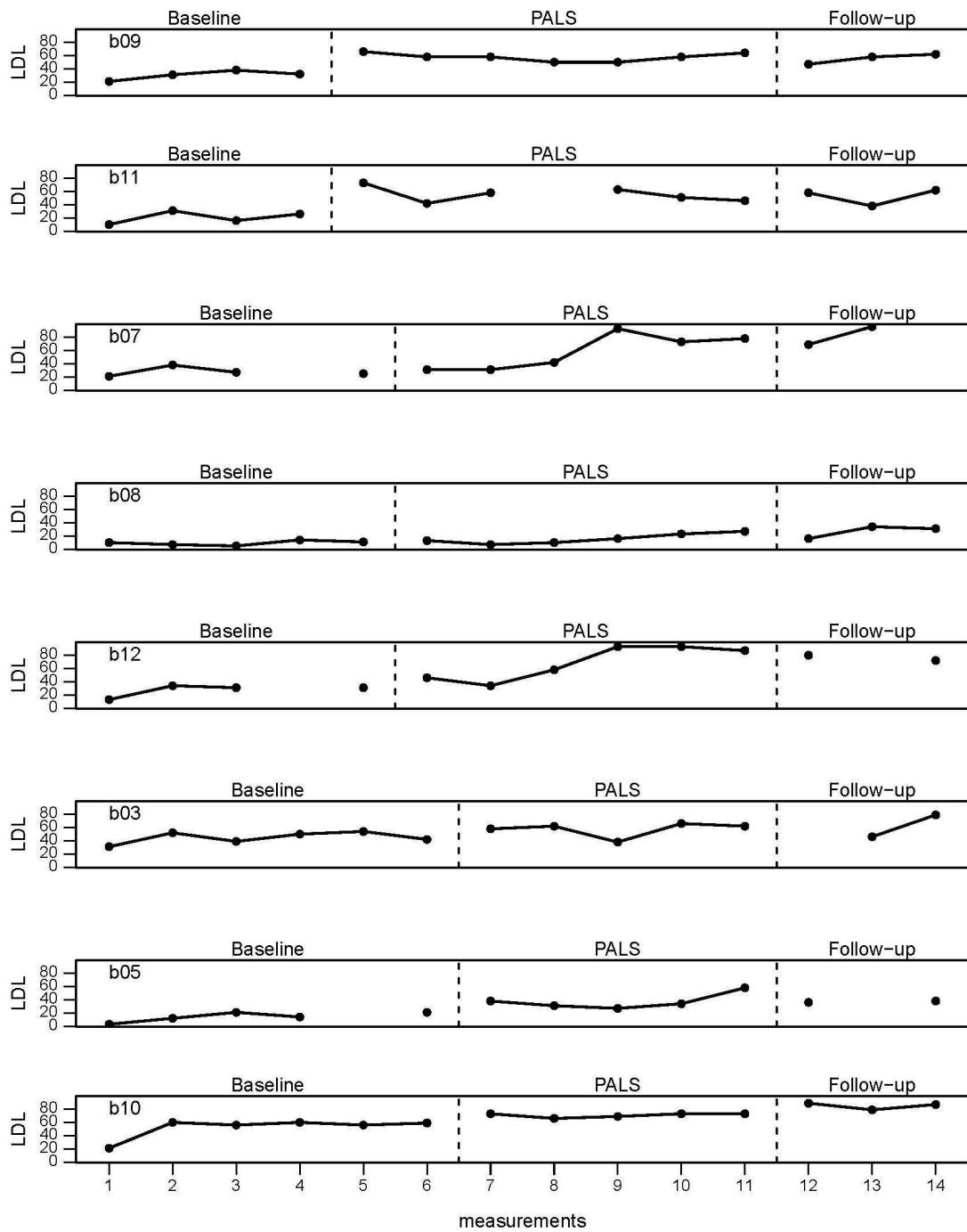
Reading comprehension

The PALS intervention significantly improved reading comprehension for most students. B09, B11, B12, B03, and B10 showed large effects from baseline to intervention, with gains maintained or increased in the follow-up phase. B07 demonstrated steady improvement with a slight increase in follow-up, indicating a moderate effect. B08 showed moderate enhancements with

less consistency, lacking significant relationships between phases. B05 showed moderate improvement, although follow-up data were unavailable. Notably, all students with behavioural problems (scores above 13 on the ITRF-G) improved in reading comprehension, with these gains sustained in the follow-up phase (see Table 3).

Table 3. Descriptive data in A, B and C phase for reading comprehension (VSL).

	N_A	N_B	N_C	M_A (SD)	M_B (SD)	M_C (SD)	MBD (AB)	Max A	Max B	Max C	NAP	Tau-U AB	Hedges' g
B09	4	7	3	9.00 (4.55)	27.29 (7.95)	22.67 (1.16)	18.29	13	42	24	100**	0.61*	2.39
B11	4	6	3	5.25 (3.40)	33.67 (10.76)	46.00 (3.61)	28.42	8	46	49	100**	0.81**	2.93
B07	4	6	2	8.00 (5.35)	30.67 (24.32)	40.00 (5.67)	22.67	13	66	44	85*	0.63*	1.05
B08	5	6	3	6.00 (6.44)	17.17 (9.72)	28.50 (7.78)	11.17	13	32	34	73	0.09	1.21
B12	4	6	2	9.00 (4.76)	27.83 (5.42)	29.50 (3.54)	18.83	12	34	32	100**	0.45	3.28
B03	6	5	3	15.16 (13.03)	61.20 (14.74)	47.33 (11.72)	46.04	34	78	56	100**	0.49*	3.05
B05	5	5	1	20.00 (11.75)	44.40 (12.60)	68.00 (N.A.)	24.4	38	60	68	90*	0.25	1.81
B10	6	5	3	18.00 (10.33)	67.40 (12.30)	71.33 (9.30)	49.4	27	84	82	100**	0.48*	4.01

Note. Measurements exclusive missing values (N); Baseline (A); Intervention (B); Follow-Up (C); Mean (M); Standard Deviation (SD); Maximum (Max); *<.05; **<.01

Figure 2. Visual Analysis in A, B and C phase for reading comprehension
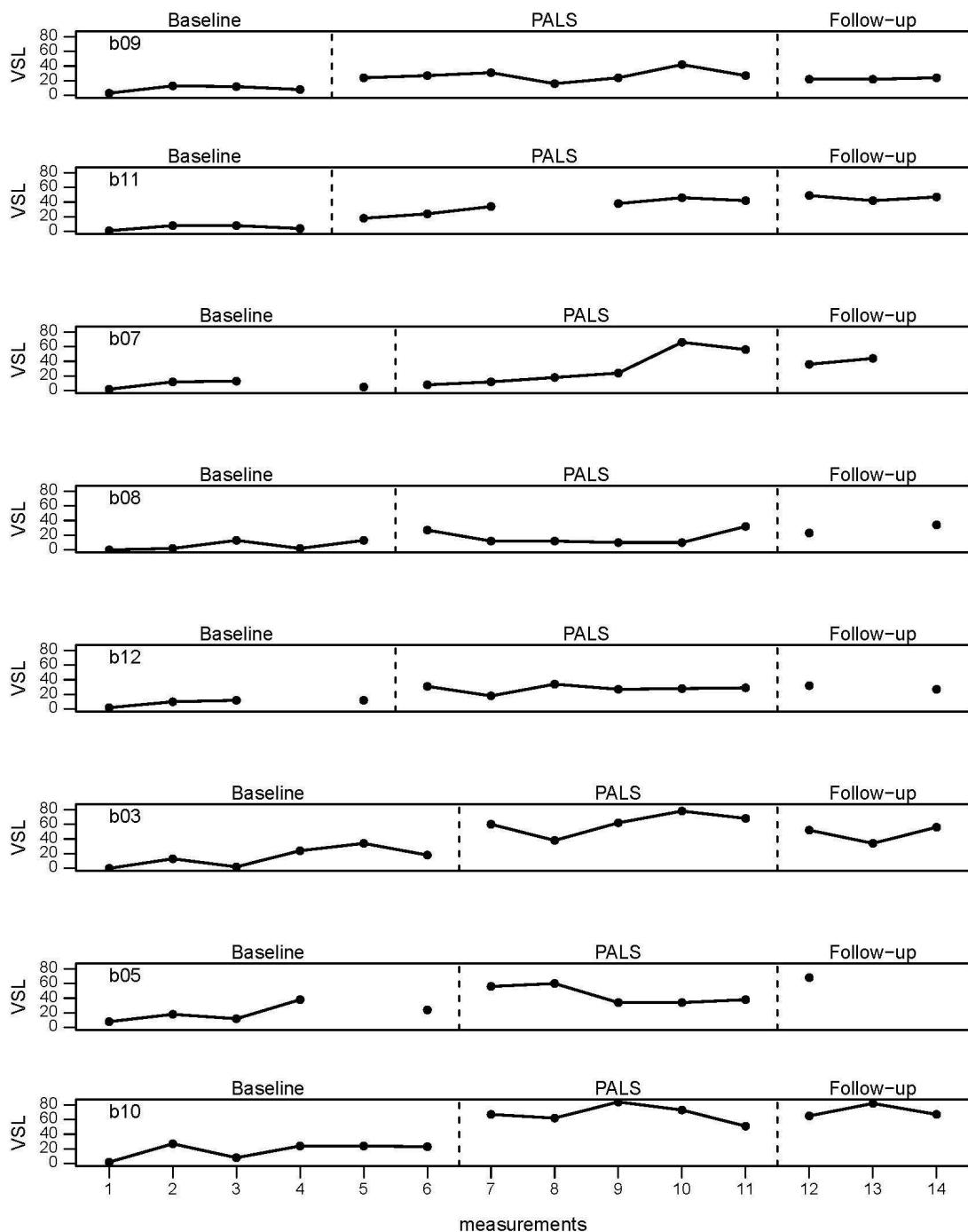
Treatment fidelity

The fidelity of the intervention was largely maintained, with high adherence to the protocol in most sessions. (1) Environmental conditions were consistently met, with unanimous positive ratings from all three raters. (2) A session plan was always available, ensuring structured sessions, (3) and all required materials were present, indicating thorough preparation and

documentation. However, while the majority of sessions were rated positively, specific deviations in (4) procedure and (5) diagnostic and feedback aspects were noted, particularly in specific sessions (1, 7, 8), but no further explanation was found in the session comments. The interrater agreement (Cooper, Heron, and Heward 2007) between the sessions was 91%.

Social validity

A total of four female (B03, B09, B11, and B12) and two male students (B05, B07) completed the social validity questionnaire. Overall, the results painted a very positive picture (mean value across all items and all students was 2.59, falling between 'somewhat true' and 'true'). Particularly, the female students agreed with the statements on average (mean value across all items was 2.97, 'true'). The two male students agreed less with the statements (mean value across all items was 1.84, between 'not true' and 'somewhat true'). Item 'I had fun correcting someone's reading' (mean value 3.33) and item 'Collecting points in the form of marbles was great' (mean value 3.5) were particularly positively rated (between 'true' and 'completely true'). It was unexpected that the students with reading and spelling difficulties who demonstrated the greatest improvement in reading fluency (B11, B07, B12) and text comprehension (B03, B10) during the intervention were not the same individuals who exhibited the highest level of agreement with social validity in the questionnaire (B11, B12: mean > 3 'agree'). On the teacher questionnaire, the overall picture was quite positive (mean value across all items was 2.5, between 'somewhat true' and 'true').

Discussion

The purpose of the study was to determine whether an adapted German version of the PALS programme improves reading fluency and comprehension skills for secondary SWRD. Overall, our findings showed noticeable improvements in reading fluency and comprehension across most students. On average, all students exhibited an improvement from the baseline to the intervention phase and subsequently scored higher on average in the follow-up phase than in the baseline phase. These findings align with previous studies (e.g. Filderman et al. 2022; Fuchs, Fuchs, and Abramson 2020; Roberts et al. 2020; Scammacca et al. 2016). The positive results are also evident, although the original PALS programme has been greatly shortened and adapted by the authors of this study. While meta-analyses suggest that long-term interventions are most effective (Hall et al. 2022; Kim et al. 2020), our short-term intervention yielded positive outcomes, thereby facilitating the transfer of knowledge into practice in comparison to the use of more time-consuming interventions.

Does an adapted German-language version of PALS improve reading fluency secondary SWRD?

Students B03, B05, and B10 displayed consistent improvements in reading fluency, with sustained or enhanced gains in the follow-up phase, indicating the intervention's effectiveness through moderate to large effects. Similarly, students B09, B11, B07, and B12 had a basic understanding of word-level comprehension (ELFE-II) and showed substantial increases, supported by large Hedge g and mostly significant relationships between phases. This suggests that the intervention was effective for these students regarding reading fluency. These findings align with Suggate (2016), who suggests that word decoding skills are a prerequisite for optimal reading interventions. However, B08 had a moderate impact of the intervention on reading fluency compared to other cases examined. He showed several variables (see Table 1; e.g. screening results) that could interfere with the successful acquisition of reading skills (e.g. Roberts et al. 2020).

Does an adapted German-language version of PALS improve reading comprehension secondary SWRD?

In terms of reading comprehension, most students showed clear improvements. This supports the idea that the adapted German-language PALS reading intervention positively affected reading comprehension, consistent with previous research (e.g. Calhoon 2005; Fuchs, Fuchs, and Abramson 2020; Harsul 2022; Völlinger, Supanc, and Brunstein 2018). Despite literature suggesting a bidirectional relationship between behaviour problems and reading difficulties (Roberts et al. 2020), four students with behavioural problems, as indicated by ITRF-G results, demonstrated improvements. The adapted PALS programme proved beneficial for these students, aligning with findings from the original PALS programme (Fuchs, Fuchs, and Abramson 2020).

How was the adapted German-language PALS intervention evaluated in terms of social validity by secondary SWRD?

Overall, the rating was positive across students and the teacher, however, there was a discrepancy between the students who provided the most positive statements and those who benefited the most from the intervention. Interestingly, the rating of the female students was more positive than that of the male students. Further evaluation might reflect inherent gender differences in responses in reading-related gender expectations (e.g. Muntoni, Wagner, and Retelsdorf 2021).

Limitations and future research

As with any study, there were limitations. Small sample size hinders generalisation, but SCR studies are effective for heterogeneous student populations and special education data collection

because they closely examine individual student learning and provide a better overview of intervention effectiveness. Additionally, COVID-19 impacted attendance and participation. Future research should include whole classes of secondary SWRD to better explore the results of German language PALS. Comparing teacher-implemented versus researcher-implemented interventions would be beneficial, as findings are mixed (Kim et al. 2020; Okkinga et al. 2018). Such comparisons could optimise PALS delivery. However, treatment fidelity and interrater reliability in this study were good. Treatment fidelity checks showed adherence to planned procedures, ensuring consistent implementation and an interrater reliability of 91% (Cooper, Heron, and Heward 2007) indicated strong agreement among raters. Validating the social validity questionnaire might provide further insights as well as having more teachers completing the form.

Conclusion

In summary, the secondary SWRD participating in this study improved their reading fluency and comprehension skills within a short intervention period through the implementation of an adapted German PALS programme. All participants exhibited improvement from the baseline to the intervention phase and were able to maintain these positive effects in the follow-up phase. This suggests that the intervention led to both short-term and sustainable improvements in reading comprehension and fluency. The positive trends shown in this pilot study are a first step to ensure the suitability of an adapted German PALS-reading programme, as recommended by Vardy et al. (2022). Overall, the results revealed that the German PALS version could be appropriate for a heterogeneous group of learners and offer practitioners a highly structured method.

References

Bakken, R. K., K.-A. B. Næss, C. J. Lemons, and H. N. Hjetland. 2021. "A Systematic Review and Meta-Analysis of Reading and Writing Interventions for Students with Disorders of Intellectual Development." *Education Sciences* 11 (10): 638. https://doi.org/10.3390/educsci11100638.

Bennett, K. J., K. S. Brown, M. Boyle, Y. Racine, and D. Offord. 2003. "Does Low Reading Achievement at School Entry Cause Conduct Problems?" *Social Science & Medicine* 56 (12): 2443–2448. https:// doi.org/10.1016/S0277-9536(02)00247-2.

Bowman-Perrott, L., H. Davis, K. J. Vannest, L. Williams, C. R. Greenwood, R. I. Parker, and C. Anderson. 2013. "Academic Benefits of Peer Tutoring: A Meta-Analytic Review of Single-Case Research." *School Psychology Review* 42 (1): 39–55. https://doi.org/10.1080/02796015.2013.12087490.

Briesch, A. M., S. M. Chafouleas, S. R. Neugebauer, and T. C. Riley-Tillman. 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.

Calhoon, M. B. 2005. "Effects of a Peer-Mediated Phonological Skill and Reading Comprehension Program on Reading Skill Acquisition for Middle School Students with Reading Disabilities." *Journal of Learning Disabilities* 38 (5): 424–433. https://doi.org/10.1177/00222194050380050501.

Calhoon, M. B., and L. S. Fuchs. 2003. "The Effects of Peer-Assisted Learning Strategies and Curriculum-Based Measurement on the Mathematics Performance of Secondary Students with Disabilities." *Remedial and Special Education* 24 (4): 235–245. https://doi.org/10.1177/ 07419325030240040601.

Calhoon, M. B., A. Sandow, and C. V. Hunter. 2010. "Reorganizing the Instructional Reading Components: Could There Be a Better Way to Design Remedial Reading Programs to Maximize Middle School Students with Reading Disabilities' Response to Treatment?" *Annals of Dyslexia* 60 (1): 57–85. https://doi.org/10.1007/s11881-009-0033-x.

Chen, M.-C., J. E. Pustejovsky, D. A. Klingbeil, and E. R. Van Norman. 2023. "Standardized Mean Differences: Flexible Methods for Single-Case Designs." *Journal of School Psychology* 98:16–38. https://doi.org/10.1016/j.jsp.2023.02.002.

Cooper, J. O., T. E. Heron, and W. L. Heward. 2007. *Applied Behavior Analysis*. 2nd ed. Upper Saddle River, NJ: Pearson.

Donegan, R. E., and J. Wanzek. 2021. "Effects of Reading Interventions Implemented for Upper Elementary Struggling Readers: A Look at Recent Research." Reading & Writing 34 (8): 1943–1977. https://doi.org/10.1007/s11145-021-10123-y.

Filderman, M. J., C. R. Austin, A. N. Boucher, K. O'Donnell, and E. A. Swanson. 2022. "A Meta-Analysis of the Effects of Reading Comprehension Interventions on the Reading Comprehension Outcomes of Struggling Readers in Third Through 12th Grades." Exceptional Children 88 (2): 163–184. https://doi.org/10.1177/00144029211050860.

Fuchs, D., E. Cho, J. R. Toste, L. S. Fuchs, J. K. Gilbert, K. L. McMaster, E. Svenson, and A. Thompson. 2021. "A Quasiexperimental Evaluation of Two Versions of First-Grade PALS: One with and One without Repeated Reading." Exceptional Children 87 (2): 141–162. https://doi.org/10.1177/ 0014402920921828.

Fuchs, D., L. S. Fuchs, and R. Abramson. 2020. "Peer-Assisted Learning Strategies (PALS): A Validated Classwide Program for Improving Reading and Mathematics Performance."

In Student Engagement, edited by A. L. Reschly, A. J. Pohl, and S. L. Christenson, 109–120. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-37285-9_6.

Fuchs, D., L. S. Fuchs, P. G. Mathes, and D. C. Simmons. 1997. "Peer-Assisted Learning Strategies: Making Classrooms More Responsive to Diversity." American Educational Research Journal 34 (1): 174–206. https://doi.org/10.3102/00028312034001174.

Fuchs, D., L. S. Fuchs, D. C. Simmons, and P. G. Mathes. 2008a. Peer Assisted Learning Strategies— Reading Methods for Grades 2-6—An Institute of Education Sciences Scaling-Up Project Teacher Manual. Nashville: Vanderbilt University.

Fuchs, D., L. S. Fuchs, D. C. Simmons, and P. G. Mathes. 2008b. Peer Assisted Learning Strategies: All-In Manual Reading Methods for Grades 2-6. Nashville: Fuchs Research Group.

Garwood, J. D., N. C. Brunsting, and L. C. Fox. 2014. "Improving Reading Comprehension and Fluency Outcomes for Adolescents with Emotional-Behavioral Disorders: Recent Research Synthesized." Remedial and Special Education 35 (3): 181–194. https://doi.org/10.1177/0741932513514856.

Gast, D. L., B. P. Lloyd, and J. R. Ledford. 2018. "Multiple Baseline and Multiple Probe Designs." In Single Case Research Methodology. Applications in Special Education and Behavioral Sciences, edited by J. R. Ledford and D. L. Gast, 239–281. 3rd ed. New York: Routledge.

Graham, S., M. Silva, and M. A. Restrepo. 2023. "Reading Intervention Research with Emergent Bilingual Students: A Meta-Analysis." Reading & Writing 36 (9): 2433–2464. https://doi.org/10. 1007/s11145-022-10399-8.

Grigoryan, A. 2020. "Major Stages of Reading Skills Development." Armenian Journal of Special Education 4 (2): 42–51. https://doi.org/10.24234/se.2020.1.1.71.

Hall, C., K. Dahl-Leonard, E. Cho, E. J. Solari, P. Capin, C. L. Conner, A. R. Henry, et al. 2022. "Forty Years of Reading Intervention Research for Elementary Students with or at Risk for Dyslexia: A Systematic Review and Meta-Analysis." Reading Research Quarterly 58 (2): 285–312. https:// doi.org/10.1002/rrq.477.

Harsul, L. K. 2022. "The Role of Peer-Assisted Learning Strategies (PALS) in Enhancing Students' Reading Skills for Second Language Acquisition." Edu-Ling: Journal of English Education and Linguistics 5 (2): 132–137. https://doi.org/10.32663/edu-ling.v5i2.2681.

Hjetland, H. N., E. I. Brinchmann, R. Scherer, C. Hulme, and M. Melby-Lervåg. 2020. "Preschool Pathways to Reading Comprehension: A Systematic Meta-Analytic Review." Educational Research Review 30:100323. https://doi.org/10.1016/j.edurev.2020.100323.

Kazdin, A. E. 2011. "Single-Case Research Designs." Child & Family Behavior Therapy 34 (1): 76–79. https://doi.org/10.1080/07317107.2012.654458.

Kim, D., Y. An, H. G. Shin, J. Lee, and S. Park. 2020. "A Meta-Analysis of Single-Subject Reading Intervention Studies for Struggling Readers: Using Improvement Rate Difference (IRD)." Heliyon 6 (11): e05024. https://doi.org/10.1016/j.heliyon.2020.e05024.

Lane, J. D., J. R. Ledford, and D. L. Gast. 2017. "Single-Case Experimental Design: Current Standards and Applications in Occupational Therapy." The American Journal of Occupational Therapy 71 (2): p71023000101–p71023000109. https://doi.org/10.5014/ajot.2017.022210.

Lee, Y., Y.-H. Cho, T. Park, and J. Choi. 2023. "Situational Interest During Individual Reading and Peer Reading Activities Using PALS: Its Relationship to Students' Reading Skills and Reading Motivations." *Educational Studies* 49 (5): 842–860. https://doi.org/10.1080/03055698.2021. 1898341.

Lenhard, W., A. Lenhard, and W. Schneider. 2020. *ELFE II. Ein Leseverständnistest für Erst-bis Siebtklässler—Version II* [ELFE II: A Reading Comprehension Test for First to Seventh Graders— Version II]. 4th ed. Göttingen: Hogrefe.

Luiselli, J. K. 2021. "Social Validity Assessment." In *Applied Behavior Analysis Treatment of Violence and Aggression in Persons with Neurodevelopmental Disabilities*, edited by J. K. Luiselli, 85–103. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-68549-2_5.

McElvany, N., R. Lorenz, A. Frey, F. Goldhammer, and A. Schilcher, and T. C. Stubbe, eds. 2023. *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre [IGLU 2021: Reading* Literacy of Primary School Children in International Comparison and Trends Over 20 Years]. Münster: Waxmann Verlag GmbH. https://doi.org/10.31244/ 9783830997009.

Moeyaert, M., D. A. Klingbeil, E. Rodabaugh, and M. Turan. 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 and Special Education* 42 (2): 94–106. https://doi.org/10.1177/0741932519855079.

Muntoni, F., J. Wagner, and J. Retelsdorf. 2021. "Beware of Stereotypes: Are Classmates' Stereotypes Associated with Students' Reading Outcomes?" *Child Development* 92 (1): 189–204. https://doi. org/10.1111/cdev.13359.

OECD. 2023. *PISA 2022 Ergebnisse: Lernstände und Bildungsgerechtigkeit* [PISA 2022 Results: Learning Levels and Educational Equity]. 1st ed. Bielefeld: wbv Media GmbH & Co. KG. https://doi.org/10. 3278/6004956w.

Okkinga, M., R. Van Steensel, A. J. S. Van Gelderen, E. Van Schooten, P. J. C. Sleegers, and L. R. Arends. 2018. "Effectiveness of Reading-Strategy Interventions in Whole Classrooms: A Meta-Analysis." *Educational Psychology Review* 30 (4): 1215–1239. https://doi.org/10.1007/s10648-018-9445-7.

Parker, R. I., K. J. Vannest, and J. L. Davis. 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.

Parker, R. I., K. J. Vannest, J. L. Davis, and S. B. Sauber. 2011. "Combining Nonoverlap and Trend for Single-Case Research: Tau-U." *Behavior Therapy* 42 (2): 284–299. https://doi.org/10.1016/j.beth. 2010.08.006.

Roberts, G. J., E. Cho, J. D. Garwood, G. H. Goble, T. Robertson, and A. Hodges. 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.

Sáenz, L., M. Lynn, S. Fuchs, and D. Fuchs. 2005. "Peer-Assisted Learning Strategies for English Language Learners with Learning Disabilities." *Exceptional Children* 71 (3): 231–247. https://doi. org/10.1177/001440290507100302.

Sanetti, L. M. H., B. G. Cook, and L. Cook. 2021. "Treatment Fidelity: What it is and Why it Matters." *Learning Disabilities Research and Practice* 36 (1): 5–11. https://doi.org/10.1111/ldrp.12238.

Scammacca, N. K., G. J. Roberts, E. Cho, K. J. Williams, G. Roberts, S. R. Vaughn, and M. Carroll. 2016. "A Century of Progress: Reading Interventions for Students in Grades 4–12, 1914–2014." *Review of Educational Research* 86 (3): 756–800. https://doi.org/10.3102/0034654316652942.

Snowling, M. J., C. Hulme, and K. Nation. 2020. "Defining and Understanding Dyslexia: Past, Present and Future." *Oxford Review of Education* 46 (4): 501–513. https://doi.org/10.1080/03054985.2020. 1765756.

Spörer, N., V. Seuring, N. Schünemann, and J. C. Brunstein. 2009. "Fostering the Reading Comprehension of Secondary School Students Through Peer-Assisted Learning: Effects on Strategy Knowledge, Strategy Use, and Task Performance." *Contemporary Educational Psychology* 34 (4): 289–297. https://doi.org/10.1016/j.cedpsych.2009.06.004.

Suggate, S. P. 2010. "Why What We Teach Depends on When: Grade and Reading Intervention Modality Moderate Effect Size." *Developmental Psychology* 46 (6): 1556–1579. https://doi.org/10. 1037/a0020612.

Suggate, S. P. 2016. "A Meta-Analysis of the Long-Term Effects of Phonemic Awareness, Phonics, Fluency, and Reading Comprehension Interventions." *Journal of Learning Disabilities* 49 (1): 77–96. https://doi.org/10.1177/0022219414528540.

Van Der Sande, L., R. Van Steensel, S. Fikrat-Wevers, and L. R. Arends. 2023. "Effectiveness of Interventions That Foster Reading Motivation: A Meta-Analysis." *Educational Psychology Review* 35 (1): 21. https://doi.org/10.1007/s10648-023-09719-3.

Vardy, E. J., S. Al Otaiba, H. L. Breadmore, S. Kung, A. Pétursdóttir, M. W. Zaru, and K. L. McMaster. 2022. "Teacher–Researcher Partnership in the Translation and Implementing of PALS (Peer- _Assisted Learning Strategies): An International Perspective." *Journal of Research in Reading* 45 (3): 517–526. https://doi.org/10.1111/1467-9817.12404.

Völlinger, V. A., M. Supanc, and J. C. Brunstein. 2018. "Examining Between-Group and Within-Group Effects of a Peer-Assisted Reading Strategies Intervention." *Psychology in the Schools* 55 (5): 573–589. https://doi.org/10.1002/pits.22121.

Volpe, R. J., G. Casale, C. Mohiyeddini, M. Grosche, T. Hennemann, A. M. Briesch, and B. Daniels. 2018. "A Universal Behavioral Screener Linked to Personalized Classroom Interventions: Psychometric Characteristics in a Large Sample of German Schoolchildren." *Journal of School Psychology* 66:25–40. https://doi.org/10.1016/j.jsp.2017.11.003.

Walter, J. 2010. *LDL Lernfortschrittsdiagnostik Lesen—Ein Curriculumbasiertes Verfahren* [LDL Reading Progress Diagnostics—A Curriculum-Based Method]. Göttingen: Hogrefe.

Wanzek, J., J. Wexler, S. Vaughn, and S. Ciullo. 2010. "Reading Interventions for Struggling Readers in the Upper Elementary Grades: A Synthesis of 20 Years of Research." *Reading & Writing* 23 (8): 889– 912. https://doi.org/10.1007/s11145-009-9179-5.

Wilbert, J., and T. Lueke. 2021. "Package 'Scan'." https://cran.r-project.org/web/packages/scan/scan. pdf.

Appendix B Article 2 *(peer-reviewed)*

Disclaimer for Dissertation Version:

This is the accepted manuscript of the article listed above. The final published version will be available from the publisher once released.

Abstract

Students with limited language proficiency often struggle with reading comprehension and fluency, which can hinder their ability to engage with academic content, ultimately impacting their overall educational achievement. Targeted reading interventions are crucial to support second language (L2) learners and ensure equal educational opportunities. Too often, these students fall so far behind that difficulties such as frequent misreading or mispronunciation of words, as well as trouble sounding out letters and blends, are systematically interpreted as signs of a reading-related learning disability. This study investigates the effectiveness of an adapted German-language Peer-Assisted Learning Strategies (PALS) program in improving the reading comprehension of L2 learners in 5th and 6th grade, using a multiple-baseline single-case design. The intervention's social validity was also assessed from the perspectives of students and teachers. Results show that reading competence improved in all participants, although progress required multiple sessions. Social validity ratings were generally positive. The study's findings are discussed with a focus on practical implications and directions for future research.

Keywords: PALS, Reading Comprehension, L2 Acquisition, Peer Tutoring, Reading Strategies, Reading-Related Learning Disabilities

Introduction

The Impact of Linguistic Barriers on Reading Proficiency

Twenty-five percent of German adolescents struggle with reading competence, with L2 learners being disproportionately affected (Organisation for Economic Co-operation and Development [OECD], 2023). Students with a migration background consistently perform below average in reading proficiency compared to their peers (OECD, 2019). These gaps reflect broader structural challenges, especially for L2 learners. PISA (Programme for International

Student Assessment [PISA]) and IQB studies (Institute for Quality Development in Education [IQB], 2022) highlight significant disparities in reading competence and German language performance among these students, partly due to linguistic barriers (IQB, 2022; OECD, 2023). Given the growing number of L2 learners, targeted interventions are essential to prevent long-term academic failure and social exclusion (Raabe, 2019). Understanding the factors that shape the reading process provides a basis for designing effective interventions. The following theoretical framework focuses on the interaction between cognitive, linguistic, and contextual elements as the foundation for this study.

Factors influencing text comprehension

The RAND Reading Study Group Model (Snow, 2002) conceptualizes reading as a multifactorial process influenced by three key components: the reader (cognitive abilities, motivation, prior knowledge), the text (genre, structure, complexity), and the reading activity (goal, context, strategies). These elements are embedded in a sociocultural context that shapes each student's reading experience.

Critical reader factors include prior knowledge, motivation, cognitive abilities, and working memory (Eason et al., 2012; Jeon & Yamashita, 2022). Text complexity and structure also influence comprehension, as more complex texts are harder to process (Pickren et al., 2022; Zhang & Lu, 2024). Successful comprehension depends on the interaction between text and reader: strong textual design combined with favorable reader conditions supports understanding, while poor design impairs it (Snow, 2002; Zhang & Lu, 2024).

Reading in Second Language Acquisition

L2 learners face greater challenges in reading than their L1 peers. Van den Bosch et al. (2020) found that the performance gap widens with age, indicating increasing disparities throughout school. Linguistic barriers such as limited vocabulary, syntax differences, and language-specific structures affect L2 learners' comprehension (Osipova & Lao, 2022; Zhang & Lu, 2024). Zhang and Zhang (2022) confirmed a strong correlation ($r = .57$) between vocabulary knowledge and comprehension. Additionally, transferring L1 rules to L2 or cultural disconnects can complicate comprehension (Bruggink et al., 2022).

Reading deficits in students with L2 are sometimes interpreted as an expression of a learning disability and diagnosed accordingly. This can be problematic, as reading difficulties may, to some extent, stem from growing up in a home where the societal language is not spoken, despite having well-developed abilities to store, process, and produce information. However, this still occurs, as it is difficult to determine when reading acquisition would have posed a problem even without bilingualism. But regardless of whether a learning disability has ever

been diagnosed, the challenges mentioned above (limited vocabulary, syntactical variations, and the complexities of language-specific structures) often lead to cognitive overload. And this, in turn, can impact other cognitive functions. For instance, Shin (2020) found a correlation ($r =$ .30) between working memory and comprehension, with stronger effects for narrative texts. Additionally, underdeveloped metacognitive skills, such as planning and monitoring, can further hinder comprehension (Vettori et al., 2024).

Intervention studies show that older students benefit more from reading strategies due to their advanced cognitive and linguistic development (Wu, 2022). Cho et al. (2021) observed a moderate effect ($d = 0.653$) of reading interventions on L2 learners' competence, with stronger effects for upper elementary students than secondary students.

Reading interventions

In Germany, the transition from elementary to secondary school (4th to 5th grade) marks a significant shift in reading demands, moving from "learning to read" to "reading to learn" (Wanzek et al., 2010). This transition requires early, targeted support to help students adapt to increased academic demands and prevent reading difficulties (Snowling et al., 2020).

To support struggling readers, particularly L2 learners, intensive and individualized interventions are essential (Filderman et al., 2022; Roberts et al., 2020). Holistic approaches that combine multiple components yield significant improvements (Suggate, 2010, 2016), focusing on strategies that guide the reading process and emphasize collaboration (Arnándiz et al., 2022; Wang & Chen, 2025). Integrating motivational elements helps reduce anxiety and fosters a more positive perception of reading (Arnándiz et al., 2022).

Since linguistic barriers and learning disabilities can overlap in their effects on reading comprehension, students with diagnosed learning disabilities also benefit from targeted, structured intervention programs (Snowling et al., 2020). These programs provide systematic support that helps address both linguistic challenges and cognitive difficulties, ultimately enhancing reading skills and overall academic success.

Peer-assisted learning strategies

Vygotsky's sociocultural theory emphasizes learning as a social process through interaction and shared experiences (Vygotsky, 1978). Peer-assisted learning reflects this perspective and has shown significant academic and social benefits (Moeyaert et al., 2021). Effective peer tutoring promotes responsibility, peer communication, and a positive error culture, enhancing student motivation (Arnándiz et al., 2022; Wang & Chen, 2025).

Duration is a critical factor in intervention success, with longer interventions showing stronger effects (Moeyaert et al., 2021; Tang et al., 2021). Among peer-tutoring models, Peer-

Assisted Learning Strategies (PALS) are particularly effective in improving reading comprehension (Fuchs et al., 1997). PALS involves four steps: Partner Reading, Retelling, Paragraph Shrinking, and Prediction Relay, fostering collaborative learning and peer support.

Numerous studies have confirmed PALS's effectiveness across age groups and learning needs, especially for struggling readers and L2 learners (Fuchs et al., 1997, 2020, 2021; Lee et al., 2023; Sáenz et al., 2005). In this respect, the PALS program offers versions tailored for reading instruction from kindergarten through high school, as well as versions for mathematics, which extend from kindergarten through grade 6 (Fuchs Research Group, 2019).

In the Kindergarten reading version (K-PALS), students practice phonological awareness, letter-sound recognition, sight word reading, and decoding, with sessions held multiple times per week. The 1st-grade version builds on these foundational skills through activities such as partner reading, word segmentation and blending, and sight word recognition, reinforcing early reading abilities. The 2nd- to 6th-grade version, which serves as the basis for this study, shifts the focus to reading fluency and comprehension, using structured, progressive strategy steps in a peer-assisted learning format. In the High School version, students further develop reading comprehension and analytical skills, working with more complex texts while continuing to engage in peer-based learning sessions (Fuchs Research Group, 2019).

In German-speaking regions, however, PALS Reading remains underexplored. Spörer et al. (2009) found positive results for seventh-grade students using German-language materials. More recently, an adapted German-language PALS program showed promising outcomes for reading comprehension and fluency in 7th grade, including benefits for L2 learners (Hertel et al., 2024).

Research Questions

Given the established effectiveness of the PALS program in English-speaking contexts, it is essential to examine its impact when adapted to a different language. Special attention should be given to L2 German students to determine whether the program benefits not only struggling readers but also those facing linguistic barriers. Investigating its effectiveness for students with specific linguistic challenges, particularly younger students transitioning to secondary school, is crucial, as they are often at an early stage of learning German as an L2.

Based on these considerations, the following research questions arise for the present study:

a) In how far does an adapted and German version of the PALS reading program enhance the reading comprehension of secondary students learning German as a second language (L2)?

b) How is an adapted and German version of the PALS reading program evaluated in terms of social validity by secondary students learning German as a second language (L2) and their teachers?

Method

Participants and Setting

Setting

The study was conducted at two secondary schools in North Rhine-Westphalia, Germany. A total of 43 students from three 5th and 6th grade classes participated in the screening phase. Teachers identified students with weak reading skills in German as an L2, with one exception: Leon, a native German speaker, was included at the teacher's request to observe whether a student without an L2 background might benefit similarly from the program.

Screenings

Prior to the study, various screenings were conducted individually or in groups to ensure appropriate participant selection. Testing took place in separate rooms to minimize distractions.

Integrated Teacher Report Form – German Language Version (ITRF-G). To gain a comprehensive understanding of the sample and form appropriate groups, the teachers of the participating students received a questionnaire. It assessed both the social and learning behaviors of the students. For this study, we utilized the short version of the German Integrated Teacher Report Form (ITRF-G, Volpe et al., 2018), which consists of 16 items. Eight items assess disruptive behavior, while the remaining eight focus on academic productivity and performance. Teachers rated each student using a scale from 0 = "not problematic" to 3 = "highly problematic". The validity of the test has been established as high according to Volpe et al. (2018).

ELFE-II - A Reading Comprehension Test for First to Seventh Graders - Version II. The ELFE-II test assesses reading comprehension in grades 1–7 at word, sentence, and text levels. For the purposes of this study, the text level was of particular importance. In word-level tasks, students pick the correct word for an image from four options. At the sentence level, students select one word from five options that best completes a given sentence. At the text level, students then read and answer a comprehension question. The test can be administered as a group assessment and takes approximately 30 minutes to complete. The validity of the test is among others evidenced by its correlation with the reading test SLS 2-9, with a correlation coefficient of $r_{ct} = .77$ (Lenhard et al., 2020).

LDL - Reading Progress Diagnostics. The LDL Test is a speed test designed to assess reading fluency. Due to the availability of 28 parallel forms, it can also be used for progress

monitoring in learning diagnostics. During the test, a student is presented with a text to read fluently within one minute. The test administrator records errors, omissions, and the total number of words read. The test is conducted individually. The validity of the LDL Test is demonstrated by correlation values of $r = .84$ with the ELFE 1-6, $r = .94$ with the SLS 1-4, and $r = .33$ with the LGVT 6-12 (Walter, 2009).

Graz Vocabulary Test (GraWo). The Graz Vocabulary Test is designed for use in elementary schools; however, it was also applied to this sample based on the students' below-average performance levels. The test measures students' receptive vocabulary and accounts for German as L2 in its norming. During the test, students hear 30 words read aloud. For each word, they are presented with four images and must select the image that best represents the word. The test can be administered as a group assessment and takes a maximum of 30 minutes to complete. Its validity has been verified through comparisons with several other related constructs (Seifert et al., 2017).

Language assessment test for children between the ages of 5 and 10 (SET 5-10). The SET 5-10 assesses expressive vocabulary in children aged 5 to 10 years. Despite a slightly older sample, this test was chosen due to expected low scores and limited alternatives for older students. The test assesses language development through subtests on vocabulary, semantics, processing speed, comprehension, production, grammar, and memory (Petermann, 2018). This study used only the first subtest, Picture Naming, to assess expressive vocabulary. In this subtest, students are individually shown images and must name them. The subtest takes a maximum of 15 minutes to complete. Correlations between all subtests of the SET 5-10 and other comparable tests range from moderate to high, indicating strong validity.

Participants

Based on the screening results, 24 students with below-average reading comprehension (VSL: percentile rank < 15; ELFE II: percentile rank < 25 at the text level) were selected for the intervention. Parental consent was obtained for all participating students. Further demographic and screening data are presented in Table 1.

During the intervention, one student withdrew due to concerns about missing regular classroom instruction. Students were excluded from the analysis if they attended fewer than three baseline sessions, which applied to three students, or missed more than four intervention sessions, which applied to a further twelve students. The basis for this decision is that if one-third or nearly one-third of all intervention sessions are missed, it can be assumed that the adapted PALS program was not sufficiently implemented to provide reliable data. Absences were due to personal absence during the school day or scheduling conflicts with other school

commitments, such as class tests or excursions. After these exclusions, the final sample con-
sisted of eight participants (N = 8).

Table 1

Demographic Characteristics of the Participants

		Leon	Jelena	Dunja	Baris	Adem	Hakim	Emre	Azad
Gender		male	female	female	male	male	male	male	male
Age		12	11	12	11	11	12	13	10
Migration		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lingua Franca German		No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
L1		German	Serbian	Serbian	Turkish	Bosnian	Arabic	Turkish	Kurdish
SEN		SED		LD			LD		
VSL		13*	1-2*	3*	8*	7*	4*	2*	4*
LDL		4-7	18-21	3-4	62	13-18	54	18	3-4
SET		6	17	1	0	2	2	0	0
GraWo		84	59	58	84	30	4	4	12
ELFE II	W	13.6	15.9	1.8	13.6	27.4	18.4	8.1	3.6
	S	1.4	18.4	1.8	9.7	9.7	1.4	9.7	2.3
	T	4.5	46	9.7	15.9	38.2	6.7	4.5	4.5
	O	2.9*	21.2	1.8*	9.7*	21.2	3.6*	4.5*	1.8*
ITRF-G		33*	17*	9	40*	14*	26*	12	30*

Note. Diagnosed Special Educational Needs (SEN); Learning Disabilities (LD); SEN in social-emotional devel-
opment (SED); VSL test reading comprehension, percentile ranks; LDL test reading fluency, percentile ranks;
SET test expressive vocabulary, percentile ranks; GraWo test receptive vocabulary, percentile ranks; ELFE II test
reading, percentile ranks at word-level (W), sentence-level (S), and text-level (T), overall results (O), *below
average reading comprehension; Integrated Teacher Report Form German (ITRF-G), *(cut-off ≥13) conspicuous
in behaviour

Interventionists and test administrators

Eight bachelor's and master's students of special education from the University of Co-
logne served as interventionists and test administrators. Each session was conducted by two
rotating interventionists, who were also responsible for baseline sessions to avoid bias. Addi-
tionally, they administered screenings and regular assessments after each session.

All interventionists received intensive training through two two-hour preparatory ses-
sions and detailed guidelines covering each phase and measurement procedure.

Design

A multiple baseline design across participants was used to attribute observed effects to the intervention, ensuring they manifested only after the intervention phase began (Ledford & Gast, 2024). This design enables the reliable measurement of irreversible behaviors, such as reading comprehension, without necessitating a return to the baseline phase. Additionally, it offers an ethical advantage by reducing the duration of the extended baseline period.

Dependent Variable

The VSL Test for assessing reading comprehension is well-suited for progress monitoring due to its 20 parallel forms. Thus, it was used not only for screening in this study but also for measuring the dependent variable of reading comprehension. In this test, students receive a booklet with two reading passages. Every seventh word is omitted and students must choose the correct word from three options. Each set includes two distractors—one linguistically, one semantically similar. As a speed test, students have four minutes to fill in as many blanks as possible. The test can be administered as a group assessment and demonstrates a parallel-test reliability of .77–.86 as well as positive findings for validity (Walter, 2013).

Intervention Material and Its Application

Two students worked together as a pair and were provided with a shared binder, referred to as their PALS folder. This folder included their team name, an organized overview of the PALS rules, and examples illustrating the implementation of individual strategy steps from the adapted PALS program. Each pair also received a book and a bookmark, which featured a summary of the general PALS behavioral guidelines. These guidelines included, among others, remaining seated quietly and working collaboratively in teams.

Each pair was given a selection of five books and was required to agree on one book to use throughout the intervention period for practicing the PALS steps. All available books were youth literature written in simple language. Before making their decision, students were presented with a brief introduction to each book. If necessary and desired, they could also make markings in their chosen book.

To further enhance motivation, each pair received an illustration of a marble jar in which they could symbolically collect marbles. These marbles were represented by emoji stickers, which the students could choose themselves after each points award. In addition to the team marble jars, there was also a points chart displayed on a Ledger-sized (11 x 17 inches) poster, representing all teams and making their scores visible to everyone. This chart was hung in the intervention room whenever possible, allowing students to view their own and other teams'

scores as needed. At the conclusion of the project, each student received a certificate of partic-
ipation for taking part in the program.

Procedures

Students were assigned to groups of four, with each group beginning the intervention at
different times. Baseline phases ranged from four to seven sessions, with randomized assign-
ment of baseline lengths. The intervention phases consisted of 16 to 18 sessions, varying based
on the length of the baseline phase. Sessions lasted approximately 30 minutes, followed by a
15-minute measurement process, and were conducted three times a week. As a result, the total
duration of the baseline and intervention phases combined was approximately eight weeks.
Three follow-up measurement sessions were conducted eight weeks after the intervention to
assess long-term effects.

Baseline

During the baseline phase, activities aimed at fostering group cohesion were conducted,
such as games and puzzles, without targeting reading skills. Interventionists followed a Base-
line Fidelity Checklist to ensure adherence to guidelines, with approximately a third of sessions
observed by an external rater. Interrater reliability was 100%.

Intervention

During the intervention phase, the adapted PALS program was gradually introduced to
students in small groups consisting of two pairs of students (tandems). The four PALS steps
were implemented sequentially: (1) Partner Reading, (2) Retelling, (3) Paragraph Shrinking,
and (4) Prediction Relay. These steps were practiced throughout the intervention period using
age-appropriate youth books.

The PALS steps began with students working in a coach-athlete structure. The "athlete"
read the text aloud and answered questions about it, while the "coach" corrected errors using
predefined error-correction procedures. Roles were switched multiple times during the PALS
sessions. Students then learned techniques for retelling the text, summarizing paragraphs, and
making predictions about how the story might continue.

Before introducing the four steps, an initial session was held to familiarize students with
the general PALS behavioral guidelines and materials. During this session, teams were formed,
ideally pairing a stronger reader with a weaker reader in line with the principles of peer tutoring
(Thurston et al., 2021). After all four steps were introduced, a guided session was conducted,
where the interventionists closely supported the students in completing all PALS steps. As stu-
dents became more proficient, the level of direct support decreased, with interventionists fo-
cusing on answering questions and providing corrections as needed.

To adapt the intervention's pacing to students' skill development, two (Partner Reading & Retelling) or three (Paragraph Shrinking & Prediction Relay) buffer sessions were available. If a buffer session was used, this was documented in the Treatment Fidelity Checklist. All students completed all four PALS steps and had at least four sessions toward the end of the intervention to independently reinforce the steps they had learned.

Throughout the intervention phase, students could earn points as teams. Points were tracked using stickers placed in a marble jar illustration and on a team scoreboard displayed in the room. Students could each award themselves one point per session if they felt they had effectively implemented the PALS steps. Additionally, the interventionist could award one point per team for good teamwork. Hence, each team could earn up to three points per session. At the end of the intervention, all students received a PALS certificate to confirm their participation in the project.

Adaptations to the original program. While the core principles of PALS - structured peer interactions, reciprocal roles, and systematic learning strategies - were maintained, the German adaptation introduced several modifications to better suit the needs of German L2 learners. This resulted in some reductions as well as linguistic and implementation-related adjustments, while the core PALS steps and behavioral guidelines remained unchanged.

However, the timeline for introducing the steps was made more flexible compared to the original program (PALS Reading for Grades 2-6). While the original program specifies 12 days for introducing the PALS steps, the adapted version allowed for a timeframe of six to 12 days, depending on how many buffer sessions were utilized. None of the small groups skipped all buffer sessions, nor did any use all available sessions, resulting in group timelines ranging from eight to 10 introductory sessions. In contrast to the original PALS program, which is designed for approximately 15 weeks with three sessions per week, the German-adapted version was significantly shortened and implemented over approximately six weeks in this study. However, the original implementation schedule of three sessions per week remained unchanged.

This flexibility not only enabled adaptive learning tailored to the students' progress but regarding implementation, also allows teachers to better integrate the PALS steps into regular classroom schedules, which often do not permit a rigid timeline for introducing specific strategies. Additionally, the preliminary practice of the PALS steps with short stories as training texts was omitted. Instead, students worked directly with the books they had selected. This approach aimed to avoid content discontinuity and maintain a focus on the PALS steps. Moreover, the points system was visually enhanced by incorporating stickers alongside numerical scores, making progress more tangible for the students.

The program was translated into German as well as linguistically adapted to align with the language and cultural context of German-speaking students. Furthermore, the terminology for "reader" and "coach" was replaced with "athlete" and "trainer", to shift the focus away from the idea that only one student per pair is responsible for reading. This modification represents a significant departure from the original program, which is designed around the peer-tutoring model that typically pairs a strong reader with a weaker reader. Since this study exclusively included struggling readers, the program was adapted because while there were slight differences in reading ability within pairs, these differences were minimal. This adjustment allowed both students in each pair to benefit from the intervention and improve their reading skills.

Fidelity

Baseline Fidelity

Interventionists followed detailed guidelines and completed a checklist after each session to document fidelity. An external rater verified approximately 40% of the sessions, with interrater reliability at 98%.

Treatment Fidelity

Similarly, fidelity during the intervention was monitored through a 21-item checklist covering environment, planning, materials, intervention procedure, and diagnostics. External observations were conducted for 40% of the sessions. The interrater reliability for these observations was 98%.

Social Validity

Students

The original PALS program has demonstrated high levels of acceptance among participating students in international studies, owing to its engaging content and step-by-step structure (Calhoon, 2005; Calhoon et al., 2003; 2010; Harsul, 2022; Lee et al., 2023; Sáenz et al., 2005). To evaluate whether the German-adapted version was also well-received and accepted by the students, its social validity was assessed at the end of the project.

Students completed a questionnaire consisting of 16 statements about their possible experiences with the adapted PALS program. Some statements focused on the students' perceived improvement in their performance, while others addressed their level of interest during the intervention. Responses were measured using a Likert scale ranging from 0 = "totally disagree" to 4 = "totally agree".

Teachers

Although the intervention was conducted by external interventionists, it is of interest to examine whether the PALS intervention also impacted regular classroom instruction and how

the students' teachers perceived the program. Teachers were given the opportunity to observe the intervention sessions if interested and were offered an introduction to the materials and the adapted program.

Similar to the students, the teachers were also provided with a questionnaire at the end of the project. The questionnaire consisted of eight statements through which the teachers could provide their assessment of the adapted program. Responses were again rated on a scale from 0 to 4. A total of two teachers participated in the survey, as the sample for this study, after being reduced due to missing data in the baseline and intervention phases, only included two different classes.

Data analysis

The entire data analysis was conducted using the Scan Package by Wilbert and Lueke (2021) for "R" For the subsequent visual analysis, graphs were created for each participant to depict reading comprehension across the baseline, intervention, and follow-up phases. These visualizations allowed for an examination of how students' reading comprehension evolved during the intervention and whether sustainable effects of the program were evident.

For the descriptive data analysis, the mean values for each phase, standard deviations, and maximum scores per phase were calculated. Furthermore, overall indices were used, such as the Non-Overlap of All Pairs (NAP) (Parker et al., 2011a) and Tau-U (Parker et al., 2011b), calculated using the formula: A vs. B + TrendB – TrendA. This allowed for a comparison of the baseline with the intervention phase while accounting for differences in trends between the two phases. Finally, a second-level regression analysis was conducted at the group level to estimate the random intercept and random slope effects between groups.

Results

Reading Comprehension

Regarding the dependent variable *reading comprehension* as measured by the VSL, the overall student responses to the intervention varied (Figure 1). However, all students exhibited an increase from Phase A to Phase B. Azad, while showing no visible improvement between these phases, demonstrated a higher mean value in the follow-up measurement compared to Phase A, indicating overall progress. For all students, it was observed that improvements during Phase B required several sessions to manifest, with occasional decreases in performance data. The follow-up measurements revealed variability across all students, although Dunja and Adem had only one and two data points, respectively, limiting the interpretability of their follow-up results.

The descriptive data (Table 2) further support these findings. A comparison of follow-up data to the baseline phase indicates that all students, except Baris, had higher overall scores in the follow-up phase. The maximum mean values per phase and per student show that peak performance predominantly occurred during Phase B. Notably, Emre achieved a significantly higher score during the follow-up phase than his highest score in Phase B, while Adem recorded the same maximum score in both Phases B and E. The mean baseline differences demonstrate percentage increases from Phase A to Phase B for all students except Azad, who showed a decline of -4.52%, a pattern consistent with visual analyses. The largest percentage increase, 62.36%, was observed for Hakim, followed by Leon with 61.16%. Adem exhibited the smallest increase at 6.71%. From Phase B to Phase E, further increases of up to 30.48% (Azad) and decreases of up to -12.93% (Leon) were noted.

Figure 1

Reading Comprehension (VSL) of Each Participant
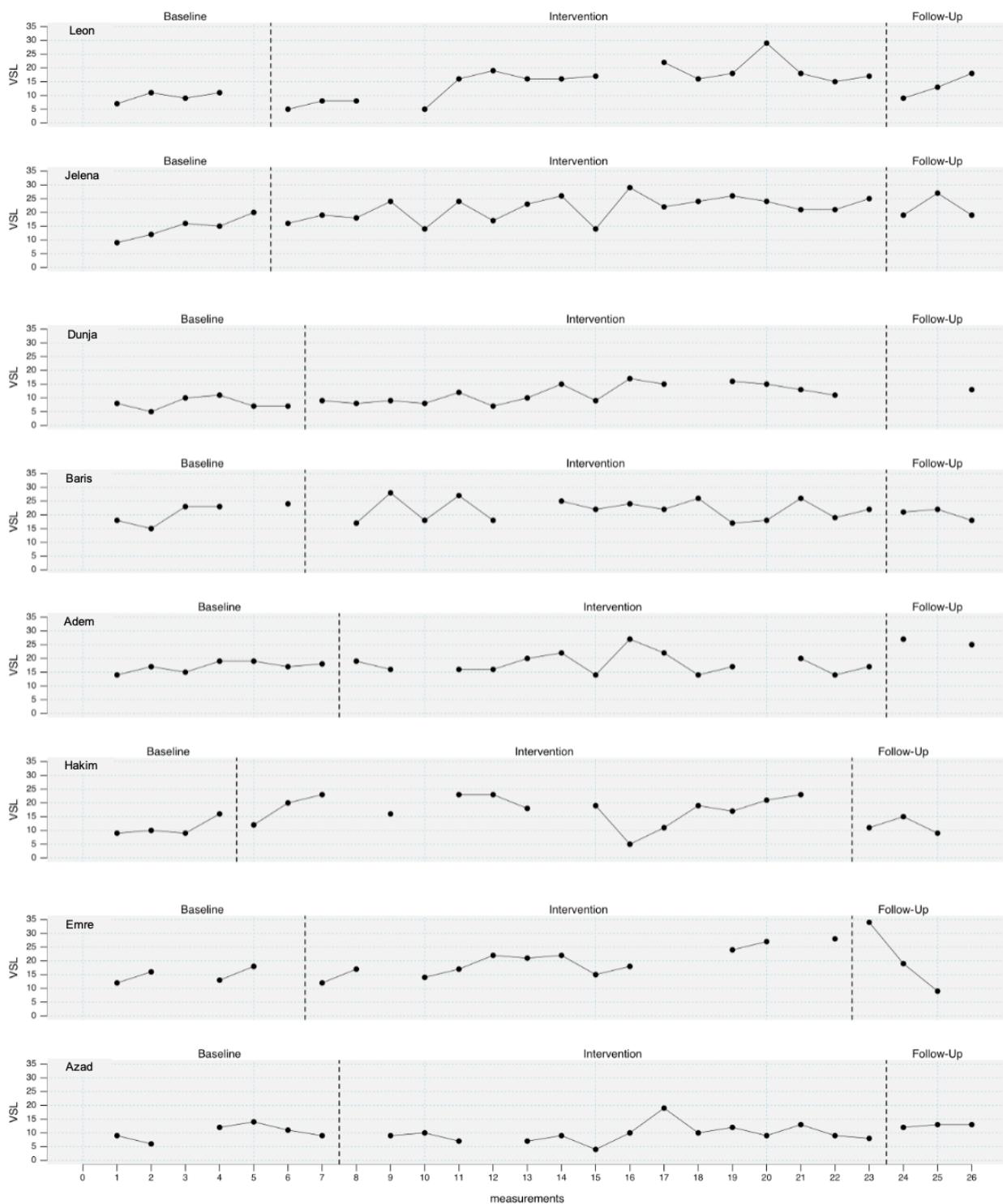
For overlap measures, no effects were detected for *tau-U* for Baris, Adem, and Azad. Medium effects were observed for the remaining students: Leon (0.47, $p < .001$), Jelena (0.42, $p < .05$), Dunja (0.48, $p < .001$), and Emre (0.58, $p < .001$). The Nonoverlap of All Pairs (NAP) analysis revealed weak effects for Baris (59.00), Adem (56.00), and Azad (42.00). Medium

effects were noted for Leon (78.00, $p < .05$), Jelena (88.00, $p < .01$), Dunja (81.00, $p < .05$), Hakim (88.00, $p < .05$), and Emre (79.00, $p = .05$).

Table 2

Descriptive Data and Overlap Indices for Reading Comprehension (VSL) of Each Participant

	N_A	N_B	N_E	M_A (SD)	M_B (SD)	M_E (SD)	Max A	Max B	Max E	NAP AB (p)	Tau-U AB (p)
Leon	4	16	3	9.50 (1.92)	15.31 (6.26)	13.33 (4.51)	11	29	18	78.00 (<.05)	0.47 (<.001)
Jelena	5	18	3	14.40 (4.16)	21.50 (4.33)	21.67 (4.62)	20	29	27	88.00 (<.01)	0.42 (<.05)
Dunja	6	15	1	8.00 (2.19)	11.60 (3.33)	13.00 (N.A.)	11	17	13	81.00 (<.05)	0.48 (<.001)
Baris	5	15	3	20.60 (3.91)	21.93 (3.90)	20.33 (2.08)	24	28	22	59.00 (.29)	0.02 (.92)
Adem	7	14	2	17.00 (1.92)	18.14 (3.76)	26.00 (1.41)	19	27	27	56.00 (.35)	0.00 (1.00)
Hakim	4	14	3	11.00 (3.37)	17.86 (5.36)	11.67 (3.06)	16	23	15	88.00 (<.05)	0.30 (.09)
Emre	4	12	3	14.75 (2.75)	19.75 (5.08)	20.67 (12.58)	18	28	34	79.00 (.05)	0.58 (<.001)
Azad	6	14	3	10.17 (2.79)	9.71 (3.45)	12.67 (0.58)	14	19	13	42.00 (.72)	-0.01 (.95)

Note. Measurements exclusive missing values (N); Baseline (A); Intervention (B); Follow-Up (E); Mean (M); Standard Deviation (SD); Maximum (Max)

Social Validity

Students

In the context of the social validity of the participants (students), the overall picture is rather positive with some deviations. For example, partner work, collecting points and choosing a book were rated positively. Overall, the students reported that they generally enjoyed the support ($M = 2.83$) and that they would like to take part again ($M = 3.33$). In addition, the students think that the PALS program could also help other students ($M = 3.00$) and that they were generally able to help each other well ($M = 2.67$). The students also showed that they liked correcting someone else and playing the role of tutor. ($M = 2.67$). It turns out that not all students understood the meaning/goal of the support well ($M = 1.67$) and that not all students liked reading in books ($M = 2.17$). There is also an unclear picture with regard to the benefit of the points

on concentrated behavior (M = 2.17). It should also be noted that Baris tended to give a lower rating for the items, which pulls down the mean value. He generally did not seem to like the support as much.

Table 3

Social Validity of Each Participant

Items	Leon	Jelena	Dunja	Baris	Adem	Hakim	Emre	Azad	Mean
We were able to help each other well with reading.	2.00	3.00	N.A.	3.00	2.00	N.A.	4.00	2.00	2.67
I think the program also helps other students with reading difficulties.	2.00	4.00	N.A.	1.00	4.00	N.A.	4.00	3.00	3.00
I understood the purpose of the program well.	2.00	0.00	N.A.	0.00	2.00	N.A.	4.00	2.00	1.67
I learned a lot during the program.	2.00	3.00	N.A.	0.00	3.00	N.A.	4.00	3.00	2.50
I enjoyed coming to the program.	2.00	2.00	N.A.	2.00	4.00	N.A.	3.00	4.00	2.83
I enjoyed the program.	2.00	2.00	N.A.	3.00	3.00	N.A.	4.00	3.00	2.83
I would take part in the program again.	2.00	3.00	N.A.	4.00	3.00	N.A.	4.00	4.00	3.33
I enjoyed reading the books.	2.00	2.00	N.A.	0.00	3.00	N.A.	3.00	3.00	2.17
Reading in pairs was great.	2.00	3.00	N.A.	4.00	3.00	N.A.	4.00	3.00	3.17
I would like to do something like this more often.	2.00	3.00	N.A.	0.00	3.00	N.A.	3.00	2.00	2.17
I had fun correcting someone's reading.	2.00	4.00	N.A.	0.00	3.00	N.A.	4.00	3.00	2.67
I enjoyed reading as a coach and as an athlete.	2.00	3.00	N.A.	0.00	2.00	N.A.	3.00	2.00	2.00
PALS in class would be great.	2.00	2.00	N.A.	0.00	3.00	N.A.	4.00	2.00	2.17
Picking out the books was great.	2.00	3.00	N.A.	4.00	2.00	N.A.	4.00	2.00	2.83
Collecting the points was great.	2.00	4.00	N.A.	3.00	3.00	N.A.	4.00	3.00	3.17
The points made me concentrate better and better.	2.00	2.00	N.A.	0.00	3.00	N.A.	4.00	2.00	2.17

Note. 0 = totally disagree; 4 = totally agree; not available (N.A.)

Teachers

With regard to social validity, which was also assessed by two teachers, it was found that there was a tendency for students to read better in class/regular lessons ($M = 2.50$) and also to show more motivation ($M = 2.50$). In addition, the teachers reported back that the students gave the impression that they always looked forward to the program ($M = 3.00$) and that the students all benefited ($M = 3.00$). Teachers differed in their opinion as to whether the students enjoyed the program ($M = 2.00$).

Table 4

Social Validity of Each Teacher

Items	Teacher 1	Teacher 2	Mean
The students also talked about the program outside of the program.	3	4	3.50
The students seemed to enjoy the program.	3	1	2.00
The program also helped the students to read correctly in class.	2	3	2.50
The students showed more motivation in class.	2	3	2.50
The students were looking forward to taking part in the program.	3	3	3.00
After the program, the students talked about it.	2	4	3.00
The students also talked about the program with people who were not involved.	1	2	1.50
I have the feeling that the students benefited from the program overall.	3	3	3.00

Note. 0 = totally disagree; 4 = totally agree

Discussion

Main Findings and Integration of Research Questions

This study aimed to explore the impact of the adapted PALS program on the reading comprehension of secondary students learning German as L2 and to examine how students and teachers evaluate the program's social validity. The findings provide initial evidence for the program's effectiveness in improving reading comprehension for L2 learners. These results offer preliminary insights while also highlighting areas for further investigation and refinement.

The results indicate that the adapted PALS program may support reading comprehension for many students, with moderate gains observed for several participants. Medium effect

sizes were identified for Leon (NAP = 78.00, Tau-U = 0.47, p < .001), Jelena (NAP = 88.00, Tau-U = 0.42, p < .05), and Emre (NAP = 79.00, Tau-U = 0.58, p < .001), suggesting that the program helped these students improve their reading skills over the intervention period. In contrast, weaker effects were noted for Baris (NAP = 59.00), Adem (NAP = 56.00), and Azad (NAP = 42.00), indicating that the program was less effective for them.

Research Question 1: To what extent does the adapted German version of the PALS program enhance reading comprehension in secondary students learning German as L2? These results suggest that the program has the potential to enhance reading comprehension for some students, although the variability in outcomes underscores the importance of addressing individual differences. According to the RAND Reading Study Group Model (Snow, 2002), reading comprehension depends on a combination of cognitive abilities, prior knowledge, and text complexity. For students with weaker effects, such as Baris and Adem, limited vocabulary knowledge may have contributed to these results, as vocabulary is closely linked to comprehension in L2 learners (Van den Bosch et al., 2020; Zhang & Zhang, 2022). This is consistent with findings from Shin (2020), who demonstrated a significant relationship (r = .30) between working memory capacity and reading comprehension in L2 learners. Cognitive overload may have further hindered progress for some students (Vettori et al., 2024).

Social validity results further enrich these findings by providing insights into how students experienced the program. Partner work (M = 3.17), point collection (M = 3.17), and book selection (M = 2.83) were rated positively by most students. Many expressed their willingness to participate again (M = 3.33) and believed the program could help other students (M = 3.00). However, some students did not fully understand the program's purpose (M = 1.67), and opinions on reading materials were mixed (M = 2.17). Baris consistently provided lower ratings, which affected the overall mean.

Research Question 2: How do secondary students learning German as L2 and their teachers evaluate the adapted German version of the PALS program in terms of social validity? Teacher feedback generally aligned with the positive student evaluations. Teachers reported that students seemed more motivated (M = 2.50) and engaged in regular classroom activities (M = 3.00). While not all students showed enthusiasm, most were described as benefiting from the program. These findings align with previous research emphasizing the motivational and collaborative benefits of peer-assisted learning (Fuchs et al., 2021; Moeyaert et al., 2021; Vygotsky, 1978). Moeyaert et al. (2021) highlight that peer tutoring can enhance both academic performance and social competence, creating a more supportive learning environment. The findings indicate that the adapted PALS program holds potential for improving reading

comprehension in L2 learners. The adapted version demonstrated that a shortened program—allowing for easier integration into the school day—still led to positive effects. Additionally, the inclusion of motivational reinforcements, such as stickers for the marble jar, supported students during the intervention. This adaptation aligns with the original program's point-based reward system, reinforcing student engagement and participation. However, the observed variability in individual outcomes suggests that several contextual and methodological factors may have influenced the results. A closer examination of these limitations is necessary to understand how they might have affected the intervention's impact and to guide future adaptations.

Limitations

Several limitations may have influenced the results. First, as a single-case research design, this study is intended to explore a specific case in depth rather than to produce broadly generalizable findings (Ledford & Gast, 2024). While the small sample size limits the direct applicability of the results to larger populations, replications of further studies on this topic can promote generalizability in the long term. The multiple-baseline design enhances internal validity, and the findings contribute new perspectives for future research. Additionally, the short duration of the intervention may have constrained long-term improvements in reading comprehension, highlighting the need for further studies to examine sustained effects over time.

According to the RAND Reading Study Group Model (Snow, 2002), reading comprehension is a multifactorial process influenced by the reader's cognitive resources, prior knowledge, and the complexity of the text. The heterogeneity of peer groups may have limited the potential for reciprocal learning, particularly for students with less developed vocabulary knowledge. Moreover, cognitive overload (Shin, 2020; Van den Bosch et al., 2020) could have further affected these students, especially when the reading materials were too complex for their current proficiency level. Research on peer-assisted learning suggests that balanced pairing of stronger and weaker readers tends to produce more robust outcomes (Fuchs et al., 1997; Sáenz et al., 2005; Thurston et al., 2021). This should be taken into account in the adapted version in the future. Fuchs et al. (1997) demonstrated the effectiveness of structured peer tutoring in improving reading fluency and comprehension among students with diverse needs. Furthermore, some reading materials were likely too complex for students with limited vocabulary, which may have reduced engagement and affected performance.

Scheduling and time constraints posed additional challenges. The intervention sessions exceeded regular class periods, making it difficult to integrate the program seamlessly into the school timetable. Irregular attendance and shortened sessions further limited the consistency and continuity of the intervention.

Implications for Future Research and Practice

Despite the methodological limitations of this study, it can nonetheless be stated that the adapted PALS program holds significant practical value for the classroom. It offers a promising approach to preventing reading delays by providing targeted support for L2 learners and enhancing reading comprehension. By promoting peer-assisted learning and addressing individual student needs, it can be effectively integrated into daily school activities to reduce the risk of reading-related learning disabilities and support continuous progress in reading skills for all students.

Regarding future research, several key recommendations emerge from the findings of this study: It is important to conduct a randomized controlled trial (RCT) to strengthen the evidence base by comparing the adapted PALS program with other reading interventions and a control group. Reducing measurement points, extending the intervention, and optimizing peer group composition could improve feasibility and effectiveness (Sáenz et al., 2005; Thurston et al., 2021). Additionally, embedding the program into regular classroom activities and aligning reading materials more closely with students' interests may foster greater engagement and long-term improvements in reading comprehension.

Furthermore, future research could explore the potential for integrating diagnostic elements into PALS, particularly to help distinguish between L2 learners with and without learning disabilities. This issue is particularly important because there is an overdiagnosis of children with L2 as LD, as the language difficulties are not related to learning a new language but to a limitation in learning. The consequences are far-reaching with children who are misdiagnosed and therefore do not receive the support they need (Dubois et al., 2020). Integrating diagnostic elements into PALS could therefore contribute to reducing the over-identification of migrant students as having learning difficulties and provide a more targeted approach to language support. Addressing these practical challenges through targeted research and practice will help ensure that the program's full potential is realized.

Conclusion

The adapted PALS program offers a promising approach for improving the reading comprehension of secondary students learning German as an L2. While the findings suggest generally positive effects, the variability in individual outcomes points to the need for continued refinement and adaptation. Future research should prioritize optimizing the program's structure and duration while exploring its implementation in diverse educational contexts. Such efforts could enhance its effectiveness and scalability, ultimately contributing to more inclusive and

evidence-based educational practices that support equitable learning opportunities for L2 learners.

References

Arnándiz, O. M., Moliner, L., & Alegre, F. (2022). When CLIL is for all: Improving learner motivation through peer-tutoring in Mathematics. *System*, *106*, article 102773. https://doi.org/10.1016/j.system.2022.102773

Bruggink, M., Swart, N., van der Lee, A., & Segers, E. (2022). Reading comprehension and multilingual students. In *Putting PIRLS to use in classrooms across the globe* (vol. 1, pp. 47–62). Springer, Cham. https://doi.org/10.1007/978-3-030-95266-2_4

Calhoon, M. B. (2005). Effects of a peer-mediated phonological skill and reading comprehension program on reading skill acquisition for middle school students with reading disabilities. *Journal of Learning Disabilities*, *38*(5), 424–433. https://doi.org/10.1177/00222194050380050501

Calhoon, M. B., & Fuchs, L. S. (2003). The effects of peer-assisted learning strategies and curriculum-based measurement on the mathematics performance of secondary students with disabilities. *Remedial and Special Education*, *24*(4), 235–245. https://doi.org/10.1177/07419325030240040601

Calhoon, M. B., Sandow, A., & Hunter, C. V. (2010). Reorganizing the instructional reading components: Could there be a better way to design remedial reading programs to maximize middle school students with reading disabilities' response to treatment? *Annals of Dyslexia*, *60*(1), 57–85. https://doi.org/10.1007/s11881-009-0033-x

Cho, Y., Kim, D., & Jeong, S. (2021). Evidence-based reading interventions for English language learners: A multilevel meta-analysis. *Heliyon*, *7*(9), article e07846. https://doi.org/10.1016/j.heliyon.2021.e07846

Dubois, P., St-Pierre, M.-C., Desmarais, C., & Guay, F. (2020). Young adults with developmental language disorder: A Systematic Review of Education, Employment, and Independent Living Outcomes. J*ournal of Speech, Language, and Hearing Research*, *63*(11), 3806–3820. https://doi.org/10.1044/2020_JSLHR-20-00127

Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader-text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology*, *104*(3), 515–528. https://doi.org/10.1037/a0027182

Filderman, M. J., Austin, C. R., Boucher, A. N., O'Donnell, K., & Swanson, E. A. (2022). A meta-analysis of the effects of reading comprehension interventions on the reading

comprehension outcomes of struggling readers in third through 12th grades. *Exceptional Children*, *88*(2), 163–184. https://doi.org/10.1177/00144029211050860

Fuchs Research Group. (2019). *What is PALS?* Vanderbilt University. https://frg.vkcsites.org/what-is-pals/

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

Fuchs, D., Fuchs, L. S., & Abramson, R. (2020). Peer-assisted learning strategies (PALS): A validated classwide program for improving reading and mathematics performance. In A. L. Reschly, A. J. Pohl, & S. L. Christenson (Eds.), *Student engagement* (pp. 109–120). Springer International Publishing. https://doi.org/10.1007/978-3-030-37285-9_6

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, *34*(1), 174–206. https://doi.org/10.3102/00028312034001174

Harsul, L. K. (2022). The role of peer-assisted learning strategies (PALS) in enhancing students' reading skills for second language acquisition. *Edu-Ling: Journal of English Education and Linguistics*, *5*(2), 133–137. https://doi.org/10.32663/edu-ling.v5i2.2681

Hertel, S., Bracht, J., Calhoon, M. B., Grünke, M., & Barwasser, A. (2024). Effects of an adapted peer-assisted learning strategies reading programme on reading fluency and reading comprehension of secondary students with or at-risk for reading disabilities. *European Journal of Special Needs Education*, 1–17. https://doi.org/10.1080/08856257.2024.2402166

Institut zur Qualitätsentwicklung im Bildungswesen [IQB]. (2022). *IQB-Bildungstrends 2021: Kompetenzen und schulische Lernbedingungen im zweiten Pandemie-Jahr* [IQB education trends 2021: Competencies and school learning conditions in the second year of the pandemic]. Waxmann.

Jeon, E., & Yamashita, J. (2022). Chapter 3. L2 reading comprehension and its correlates: An updated meta-analysis. In E. Jeon & Y. In'nami (Eds.), *Understanding L2 Proficiency: Theoretical and meta-analytic investigations* (pp. 29–86). John Benjamins Publishing Company. https://doi.org/10.1075/bpa.13.03jeo

Ledford, J. R., & Gast, D. L. (Eds.). (2024). *Single case research methodology: Applications in special education and behavioral sciences* (4th ed.). Routledge. https://doi.org/10.4324/9781003294726

Lee, Y., Cho, Y.-H., Park, T., & Choi, J. (2023). Situational interest during individual reading and peer reading activities using PALS: Its relationship to students' reading skills and reading motivations. *Educational Studies*, *49*(5), 842–860. https://doi.org/10.1080/03055698.2021.1898341

Lenhard, W., Lenhard, A., & Schneider, W. (2020). *ELFE II: Ein Leseverständnistest für Erst-bis Siebtklässler – Version II* [ELFE II: A reading comprehension test for first to seventh graders - Version II] (4th ed.). Hogrefe.

Moeyaert, M., Klingbeil, D. A., Rodabaugh, E., & 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 and Special Education*, *42*(2), 94–106. https://doi.org/10.1177/0741932519855079

Organisation for Economic Co-operation and Development [OECD] (Ed.). (2023). *PISA 2022 Ergebnisse: Lernstände und Bildungsgerechtigkeit* [PISA 2022 Results: Learning Levels and Educational Equity]. Waxmann. https://doi.org/10.3278/6004956w

Organisation for Economic Co-operation and Development [OECD] (Ed.). (2019). *PISA 2018 Ergebnisse: Lernstände und Bildungsgerechtigkeit* [PISA 2018 Results: Learning Levels and Educational Equity]. Waxmann. https://doi.org/10.31244/9783830991007

Osipova, A. V., & Lao, R. S. (2022). Breaking the cycle of failure for culturally and linguistically diverse learners with exceptional needs: Recommendations for improvement of teacher preparation programs. *New Waves-Educational Research and Development Journal*, *25*(2), 1–25.

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

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011b). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, *42*(2), 284–299. https://doi.org/10.1016/j.beth.2010.08.006

Petermann, F. (2018). *Sprachstandserhebungstest für Kinder im Alter zwischen 5 und 10 Jahren* [Language assessment test for children between the ages of 5 and 10] (3rd ed.). Hogrefe.

Pickren, S. E., Stacy, M., Del Tufo, S. N., Spencer, M., & Cutting, L. E. (2022). The contribution of text characteristics to reading comprehension: Investigating the influence of text emotionality. *Reading Research Quarterly*, *57*(2), 649–667. https://doi.org/10.1002/rrq.431

Raabe, I. J. (2019). Social exclusion and school achievement: Children of immigrants and children of natives in three European countries. *Child Indicators Research*, *12*(3), 1003–1022. https://doi.org/10.1007/s12187-018-9565-0

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

Sáenz, L. M., Fuchs, L. S., & Fuchs, D. (2005). Peer-assisted learning strategies for English language learners with learning disabilities. *Exceptional Children*, *71*(3), 231–247. https://doi.org/10.1177/001440290507100302

Seifert, S., Paleczek, L., Schwab, S., & Gasteiger-Klicpera, B. (2017). *Grazer Wortschatztest* [Graz vocabulary test]. Hogrefe.

Shin, J. (2020). A meta-analysis of the relationship between working memory and second language reading comprehension: Does task type matter? *Applied Psycholinguistics*, *41*(4), 873–900. https://doi.org/10.1017/S0142716420000272

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension.* RAND Corporation. https://www.rand.org/pubs/monograph_reports/MR1465.html

Snowling, M. J., Hulme, C., & Nation, K. (2020). Defining and understanding dyslexia: Past, present and future. *Oxford Review of Education*, *46*(4), 501–513. https://doi.org/10.1080/03054985.2020.1765756

Spörer, N., Seuring, V., Schünemann, N., & Brunstein, J. C. (2009). Fostering the reading comprehension of secondary school students through peer-assisted learning: Effects on strategy knowledge, strategy use, and task performance. *Contemporary Educational Psychology*, *34*(4), 289–297. https://doi.org/10.1016/j.cedpsych.2009.06.004

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

Suggate, S. P. (2010). Why what we teach depends on when: Grade and reading intervention modality moderate effect size. *Developmental Psychology*, *46*(6), 1556–1579. https://doi.org/10.1037/a0020612

Tang, S., Irby, B. J., Tong, F., & Lara-Alecio, R. (2021). The effects of cooperative, collaborative, and peer-tutoring strategies on English learners' reading and speaking proficiencies

in an English-medium context: A research synthesis. *SAGE Open*, *11*(4), 21582440211060823. https://doi.org/10.1177/21582440211060823

Thurston, A., Cockerill, M., & Chiang, T.-H. (2021). Assessing the differential effects of peer tutoring for tutors and tutees. *Education Sciences*, *11*(3), 97. https://doi.org/10.3390/educsci11030097

Van den Bosch, L. J., Segers, E., & Verhoeven, L. (2020). First and second language vocabulary affect early second language reading comprehension development. *Journal of Research in Reading*, *43*(3), 290–308. https://doi.org/10.1111/1467-9817.12292

Vettori, G., Casado Ledesma, L., Tesone, S., & Tarchi, C. (2024). Key language, cognitive and higher-order skills for L2 reading comprehension of expository texts in English as foreign language students: A systematic review. *Reading and Writing*, *37*(9), 2481–2519. https://doi.org/10.1007/s11145-023-10479-3

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 of German schoolchildren. *Journal of School Psychology*, *66*, 25–40. https://doi.org/10.1016/j.jsp.2017.11.003

Vygotsky, L. S. (1978). Interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind and society: The development of higher psychological processes* (pp. 79–91). Harvard University Press.

Walter, J. (2009). *LDL Lernfortschrittsdiagnostik Lesen – Ein curriculumbasiertes Verfahren* [LDL Reading Progress Diagnostics – A Curriculum-Based Method]. Hogrefe.

Walter, J. (2013). *VSL: Verlaufsdiagnostik sinnerfassenden Lesens* [VSL: Progress diagnostics for reading comprehension]. Hogrefe.

Wang, L., & Chen, C. J. (2025). Factors affecting student academic performance: A systematic review. *International Journal on Studies in Education*, *7*(1), 1–47. https://doi.org/10.46328/ijonse.276

Wanzek, J., Wexler, J., Vaughn, S., & Ciullo, S. (2010). Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. *Reading & Writing*, *23*(8), 889–912. https://doi.org/10.1007/s11145-009-9179-5

Wilbert, J., & Lueke, T. (2021). *Package 'scan'*. https://cran.r-project.org/web/packages/scan/scan.pdf

Wu, H. (2022). A brief review of the effect of reading strategy instruction on L2 reading comprehension: A discussion of types, amount, L2 proficiency, and age. *Journal of Language Teaching*, *2*(3), 1–5. https://doi.org/10.54475/jlt.2022.003

Zhang, X., & Lu, X. (2024). Testing the relationship of linguistic complexity to second language learners' perceptions of text difficulty. *Language Learning*, *74*(3), 672–706. https://doi.org/10.1111/lang.12633

Zhang, S., & Zhang, X. (2022). The relationship between vocabulary knowledge and L2 reading/listening comprehension: A meta-analysis. *Language Teaching Research*, *26*(4), 696–725. https://doi.org/10.1177/1362168820913998

Appendix C Article 3 *(peer-reviewed)*

Disclaimer for Dissertation Version:

Abstract

Reading and writing are enormously valuable, influencing every area of life beyond school. However, reading and writing difficulties are on the rise, especially among second-language students. The primary objective of this study is to design a classroom intervention that simultaneously promotes reading and writing and is easy to implement. In a pilot study, we implemented a story-map reading intervention using the SRSD model to promote reading comprehension and text production in third and fourth-grade elementary students with reading difficulties (N = 60), several of them with German as a second language (*n* = 15). We implemented an experimental design with an intervention group and a control group with pretest, posttest, and follow-up measurements. The results of the ANCOVA demonstrate that the intervention effectively promotes sentence and text comprehension in a very short time with significant key effects among the group (sentence comprehension, $\eta2 = 0.17$; text comprehension, $\eta2 = 0.36$) and stable follow-up data. We observed only a descriptive improvement in text writing quality; the results of the ANCOVA are not meaningful due to violations. Moreover, within the framework of social validity, the data displayed predominantly positive feedback on the intervention. This study provides a promising direction for using a story map as an instructional tool to enhance students' reading comprehension and possibly text writing quality.

Keywords: Reading Intervention, Reading and Writing Ability, Story-Map, Selfregulated Strategy Development

Introduction

Reading and its significance

The ability to read is one of the most valuable skills in society. However, 10% to 20% of the population experience reading difficulties (IDA, 2020), and 25.4% of fourth-grade students do not reach the reading standard (McElvany et al., 2021). Furthermore, 1 in 3 students in Germany have a migration background, often with German as an L2. This group of students can face significant challenges regarding reading and writing (Han & Hiver, 2018), which can lead to limited education, increasing the risk of unemployment and economic disadvantage (Macdonald et al., 2016). The Programme for International Student Assessment (PISA) revealed that in 2018 and 2022, German students' reading performance had regressed compared to the previous years, requiring further support, especially in reading comprehension (Organisation for Economic Cooperation & Development, 2019), a specific key skill beyond school (Oakhill et al., 2019). Given the critical role of reading proficiency and literacy, this study aims to address the persistent challenges faced by German L2 learners, focusing on interventions that improve sentence reading, text comprehension, and text writing to provide a comprehensive literacy support system.

Challenges in reading comprehension

Reading comprehension difficulties often only become apparent as texts become more complex (e.g., Catts et al., 2012). Reading comprehension is a multidimensional construct that encompasses not only the ability to recognize words but also the integration and interpretation of information (Oakhill et al., 2019). It is crucial to be able to make inferences to sufficiently understand the content of a text. Difficulties in reading comprehension frequently stem from limited memory capacity; thus, students may not always be able to draw inferences from what they read (Cain, 2022; Shin et al., 2020, for L2 students). Furthermore, enhancing sentence-level fluencyis crucial for L2 learners to not only process individual sentences but also build a framework for understanding larger text structures. Interventions that simultaneously enhance the automaticity of sentence processing and the strategic understanding of texts can bridge the gap between these fundamental skills and higher-order comprehension abilities. Specifically, students reading and writing in their L2 have greater hurdles compared to their L1 peers, predominantly regarding word reading and reading comprehension, which can cause difficulties in summarizing a text (Marzec-Stawiarska, 2016).

Low reading and writing skills can be due to specific challenges in second language acquisition. In particular, L2 learners often encounter literacy difficulties concerning vocabulary and syntactic structures, which can impede both reading and writing (Grabe, 2016). Of

concern is a potential masking effect, where weak reading skills, exacerbated by L2 hurdles, obscure additional individual needs such as dyslexia or similar conditions (Zhang & Wang, 2023). Furthermore, L2 readers often rely on background knowledge, which could reveal notable differences compared to their experiences in their native language (L1), as their understanding of basic social and cognitive constructs can differ (Grabe, 2016). Gunderson and D'Silva (2016) assert that L2 outnumber L1 learners and have unique requirements, emphasizing the necessity for targeted interventions to alleviate cognitive load and integrate linguistic and content-related aspects through visual aids and strategic support (Grabe, 2016).

The relationship between reading and writing

Writing is a crucial communication skill and is interconnected with reading. Graham and Hebert (2011) outlined the reading and writing connection and argued that students frequently needed specific support in reading and writing simultaneously. In this context, Graham et al.'s (2017) meta-analysis also demonstrates the relationship between reading and writing. The authors demonstrated that reading interventions could improve students' writing performance by strengthening writing skills with significant effects on writing in general (effect size [ES] = 0.57). Furthermore, the effects of reading instruction on writing were maintained over time. Consequently, the combined promotion of reading and writing appears to offer promising results.

Combining effective intervention components into a single intervention

Reading comprehension interventions

Al Otaiba et al. (2023) found an overall effect size for reading interventions in elementary grades (0.39). In the upper primary grades, interventions that included multiple components (addressing more than one reading dimension) had stronger effects (0.95). Furthermore, explicit and systematic instruction was positively correlated with reading improvement in upper elementary grades. The authors advocate implementing balanced reading and writing support in classrooms.

Several meta-analyses and reviews have focused on fostering reading comprehension with several key findings. Cho et al.'s (2021) meta-analysis of reading interventions revealed an overall mean effect of d = 0.65 among K-12 L2 learners, indicating that appropriate interventions for L2 learners positively influenced reading proficiency. Medium-size groups and strategy interventions encompassing the activation of prior knowledge and expression of visualizations were found to be more effective than basic reading skills training. Moreover, upper elementary students exhibited higher effect sizes compared to secondary students, with interventions targeting reading comprehension demonstrating particularly pronounced effects.

Consequently, this finding highlights the significance of interventions implemented during the elementary school years. The authors also advocate a multitiered reading system for 5–15 individuals, as this approach yielded more substantial effects than in small-group instruction. However, further research is needed to examine the impacts on L2 students.

Based on previous findings on reading strategy instruction, Okkinga et al. (2018) focused on whole-classroom interventions and found smaller effects than in other studies of smaller-group instruction. However, the authors note that strategy instruction seems to be durable and clarify that teaching reading strategies in the classroom can influence students' reading comprehension when initially implemented in the third grade. Sencibaugh (2007) found that visually dependent strategies (0.94) and text structure-based strategies (1.18) appeared to be particularly effective. A graphic organizer is an example of an organizational strategy focusing on structure.

Organizing information in a text serves as a tool for conveying text structures and presenting information from a text visually in maps and diagrams. Sun et al.'s (2021) meta-analysis of the relationship between reading strategies and reading comprehension revealed that organizational strategies positively influenced reading comprehension in upper elementary grades. Furthermore, using graphic organizers positively moderated the effectiveness. Hence, text-structure instruction should be part of the primary school curriculum.

One can teach a strategy using the self-regulated strategy development (SRSD) model by Harris and Graham (1996). The model consists of six phases (develop background knowledge, discuss it, model it, memorize it, and support it) helping students apply a strategy themselves. Using the SRSD model to teach reading strategies can produce moderate to high effects (Gillespie Rouse & Sandoval, 2018). Initial studies on reading have also confirmed the effectiveness of SRSD on students' reading comprehension (Sanders, 2020; Sanders et al., 2019). Furthermore, Chen et al. (2021) compared the impact of non-SRSD and SRSD instruction and found that the latter was more effective in the context of L2 students also employing peer tutoring.

Reading strategies can be particularly effective when learned in a peer-supported environment. One way to create a peer-supported environment is through peer tutoring, where two students work on the tasks cooperatively while helping and learning from each other. Several studies on L1 and L2 learners (e.g., Bowman-Perrott et al., 2013; Bruggink et al., 2022; Tang et al., 2021) have shown that employing peer tutoring as an additional component to increase motivation can be effective.

The story map as a graphic organizer

The story map is a frequently used option for a graphic organizer and a specific application of the above components of facilitation, primarily combining structurebased and organizational strategies. This method organizes content from a narrative text using a visual template to increase the reading comprehension and attention of students with and without disabilities. One approach to addressing the existing issues of fundamental reading difficulties among students even in upper elementary grades is to focus on narrative rather than informative texts. While expository texts primarily provide detailed background knowledge within a specific field, narrative texts facilitate the integration of students' own experiences with the content of the story, fostering a holistic understanding of connections, perspectives, and constellations (Bogaerds-Hazenberg et al., 2020). Furthermore, at higher grade levels, narrative texts demand essential cognitive skills such as inference, prediction, and evaluation (Clinton et al., 2020). Research has shown that fostering students' reading comprehension using narratives is also effective in improving L2 literacy (e.g., Bruggink et al., 2022), even though little research on this approach exists in the L2 area to date. In a recent meta-analysis, Cure et al. (2020) confirmed the effectiveness of the story-map strategy among students with different disabilities, especially students with learning disabilities in different grades (Ciullo & Reutebuch, 2013; Onachukwu et al., 2007; Stagliano & Boon, 2009). Furthermore, Alves et al. (2015) demonstrated the long-term effects of the story map on reading comprehension among students with reading difficulties. Teaching this strategy typically features direct instruction, modelling the steps, and ongoing practice (Cure et al., 2020).

Adding motivational components to an intervention package

McBreen and Savage's (2020) systematic review of the effects of motivational reading instruction on students' reading competency and enthusiasm revealed the following results, independent of grade level and group size: reading comprehension ($g = 0.40$), reading competency ($g = 0.13$), and motivation ($g = 0.43$). In terms of learning in various domains, motivational amplifiers also play a crucial role in students' academic performance. For example, positive reinforcement for students (McLeskey et al., 2017) can reduce potential anxiety and increase their motivation and willingness to exert themselves. However, teachers should consider the duration of the intervention to maintain students' willingness and focus. Shorter interventions (two months or less, $\beta = 0.530$) are more effective than longer interventions ($\beta = -0.454$; Sohn et al., 2023).

Research aim

On the basis that it is vital to encourage students' reading and writing to provide equal educational opportunities and promote writing through a reading intervention, we design a

multicomponent intervention that primarily targets reading and comprises several components that research has demonstrated to be highly effective. We employ the graphic organizer as the strategy tool with the help of SRSD, peer tutoring, and motivational components.

We conduct the intervention with an entire class to make it easier for teachers to implement it in the classroom and target a diverse student population with reading difficulties. The intervention aims to address the following research questions:

(1) Does a combined reading intervention influence sentence reading among students with reading difficulties with and without German as an L2?

(2) Does a combined reading intervention influence text reading among students with reading difficulties with and without German as an L2?

(3) Does a combined reading intervention influence the text writing of students with reading difficulties with and without German as an L2?

(4) How is the combined reading intervention evaluated in the context of social validity by the persons involved?

Methods

Study participants and setting

The study took place at a low socioeconomic urban elementary school in North-Rhine Westphalia, Germany and focused on the third and fourth grades (ages 9–10). Before the study commenced, we obtained consent from the legal guardians and the participants. We assigned the participants pseudonyms for data protection reasons. Four third-grade classes participated in the study, of which two functioned as experimental groups and two as controls. We collected general data from the children via a teacher's questionnaire asking about age, gender, special needs, and L2. In addition, we employed a German reading screening (ELFE; Lenhard et al., 2020) and a writing quality test to capture the dependent variables, which are described in the assessment section.

In total, 60 children participated in the study, all of whom exhibited reading difficulties with a percentile rank of < 30 in the ELFE screening. The remaining children with average reading ability participated in the intervention. However, we collected no data for this group (to carry out the intervention in the regular classroom and not exclude the other children). A percentile rank between 15–30 included all children at risk of reading difficulties. We randomly assigned the four classes to either an experimental or a control condition. We surveyed 13 students from Class 3a and 17 students from Class 3c as part of the experimental group and 14 students from Class 3b and 16 students from Class 3d as a control. A T-test revealed that the groups were similar in many characteristics. The treatment group contained significantly more

children with an L2 and slightly more difficulties in text reading. All the L2 children had started learning German at age 3.

Research design and dependent variables

Research design

We employed an experimental design with an experimental group and a control group involving pretest, posttest, and follow-up measurements. The treatment group received a story-map reading intervention, and the control group attended a regular class without specific reading training. We randomly assigned the participants to one of the conditions. Eight master's students with special needs education received intensive training in measuring the dependent variables and implementing the treatment. Four were responsible for delivering the treatment, while two were always present in the classroom. The four interventionists rotated during the weeks. Two were responsible for data collection and assessing text writing quality, and the other two were responsible for ensuring treatment fidelity and observing the treatment group and control groups. The intervention took place for one hour three times a week over four weeks, resulting in 12 intervention sessions that all children from the treatment group attended. The intervention started one week after the screenings and pretests. We conducted a posttest the week after the end of the intervention. We took the follow-up measurement two months after the intervention ended, including two weeks' holiday. We allocated the students to control and treatment classes, by randomly assigning the students in a class (e.g., Class 3d) to either control or treatment conditions, resulting in two classes: treatment and control. These classes did not include the original students because of the randomization.

Pretest, posttest, and follow-up measurements

We expected the treatment to influence the three dependent variables, which are described in detail in the following sections. The first dependent variable was reading at the sentence level, and the second was reading at the text level. We used the percentile rank for both variables. The third dependent variable was text writing quality since reading interventions can also influence writing (Graham et al., 2017).

Reading test (ELFE II). ELFE II is a standardized reading comprehension test for first to seventh-graders that measures reading fluency, accuracy, and comprehension at the word, sentence, and text levels (Lenhard et al., 2020). The sentence comprehension sub-test presents sentences for which a suitable word is selected from five alternatives and inserted into the sentence at a certain point. The text comprehension subtest presents short texts with one or more related questions. Each item consists of one question and four corresponding answer alternatives, from which participants must select an appropriate response. The entire class took part in

the ELFE II as a group exercise. The retest reliability for this test procedure ranged from rtt = 0.81 to rtt = 0.90 for the subtests and rtt = 0.93 for the total score (after 30 days). Parallel test reliability was rtt = 0.93. We used the respective percentile ranks for the data analysis. The components of the ELFE test, particularly sentence and text comprehension, are critical indicators for assessing reading comprehension (Table 1).

Table 1

Participant information for control and treatment groups

Variable	N	Control (N = 30) M(SD)	N	Treatment (N = 30) M(SD)	*p*
Gender	m = 14 f = 16	1.53 (.51)	m = 19 f = 11	1.37 (.49)	0.18
Age		9.57 (.94)		9.73 (.91)	0.66
Special needs	1	–	2	–	0.85
German L2	15	0.50 (5.1)	23	0.77 (.43)	< 0.05
Reading (text)		19.97 (11.22)		16.97 (10.15)	< 0.01
Reading (sentence)		24.02 (6.47)		20.11 (9.23)	0.19
Text writing quality		6.76 (3.80)		7.00 (3.40)	0.70

Percentile Rank (PR); Second Language (L2); Special Needs (SN); Mean (M); Standard Deviation (SD); male (m); female (f)

Text writing quality. The children wrote stories based on writing prompts given during the screening and as a dependent variable measure. For each story, the children selected one of three writing prompts; the writing prompts were adapted to the student's interests (e.g., alone at home in the evening). The selected writing prompts were not repeated. The interventionists provided the children with lined paper to handwrite their stories. We determined text writing quality using a six-point Likert-scale text writing quality rubric based on the Teacher Evaluation of Story Elements (TESE) scale by Troia and Graham (2002). This rubric comprises five categories to determine whether certain elements occur in a story: 1) setting (when, where, who), 2) problem, 3) actions to solve the problem, 4) consequences of the actions, and 5) emotions of the characters. The first element contains an assessment of the setting description: the story timeframe (when), the locations (where), and the characters involved (who). The second

element assesses the extent to which the story presents a problem to solve (problem). Element 3 follows from the second criterion by assessing the described actions that contribute to solving the problem (actions to solve the problem). Element 4 is also directly related to the previous assessments, as it focuses on the consequences of the actions described and how the story ends (consequences of the actions). Element 5 offers the opportunity to assess the described emotions of the characters involved (character emotions). We used raw scores for the analysis (Table 2).

Table 2

Descriptive statistics for reading comprehension and text writing quality

Measures	Pretest *M*(*SD*)	Posttest *M*(*SD*)	Follow-Up *M*(*SD*)
Reading (sentence)			
Control (N = 30)	24.02 (8.47)	23.54 (11.25)	24.39 (10.24)
Treatment (N = 30)	20.11 (11.24)	35.12 (9.55)	37.86 (10.35)
Reading (text)			
Control (N = 30)	19.97 (12.21)	19.03 (10.29)	17.65 (11.03)
Treatment (N = 30)	16.97 (10.15)	39.11 (11.01)	37.80 (12.63)
Text writing quality			
Control (N = 30)	7.19 (3.88)	5.96 (2.58)	7.06 (3.25)
Treatment (N = 30)	6.45 (3.08)	8.34 (3.37)	9.14 (3.62)

Means (M), Standard Deviations (SD); percentile ranks are shown for reading (<31 underdeveloped reading performance) and raw scores for text writing quality

Students could earn a maximum of five points for each of the categories, ranging from 0 (no aspect found) to 5 (all aspects found) with 5 indicating the highest quality. Therefore, each composition could earn between 0 and 25 points. Assessing writing progress typically involves the use of rubrics created by researchers. This approach is common because no better alternatives exist. The reliability of these rubrics is sometimes questionable (see Rezaei & Lovorn, 2010). In our sample, Cronbach's alpha was 0.83 for the five items. The appendix (Table A3) gives the correlations between all the scores within the condition. Two assessors independently rated each composition, taking the mean score as the dependent variable. The IoA for text writing quality is 96.44%, which we consider reliable, as the agreement exceeds 80% (Hausman et al., 2022).

Intervention material

As a central component of the intervention, each child received a *My Reading Folder* comprising 12 short stories. For each story, the children had to complete a story map printed below the stories. The story map contained various fields for story-related questions: Who?

When? Where? What was the problem? How was the problem solved? The goal was to complete the individual fields after reading the stories, based on information from the texts, to be able to better extract and organize the information in the stories. A matching symbol accompanied each of the storymap questions (Fig. 1), offering the students visual aids in addition to the written words. The children also received individual story-map questions on small cards so that they could recall the points of the story map at any time while reading the stories. The story cards were also available as large copies for visualization on the board in the classroom.

Figure 1

Story-map organizer
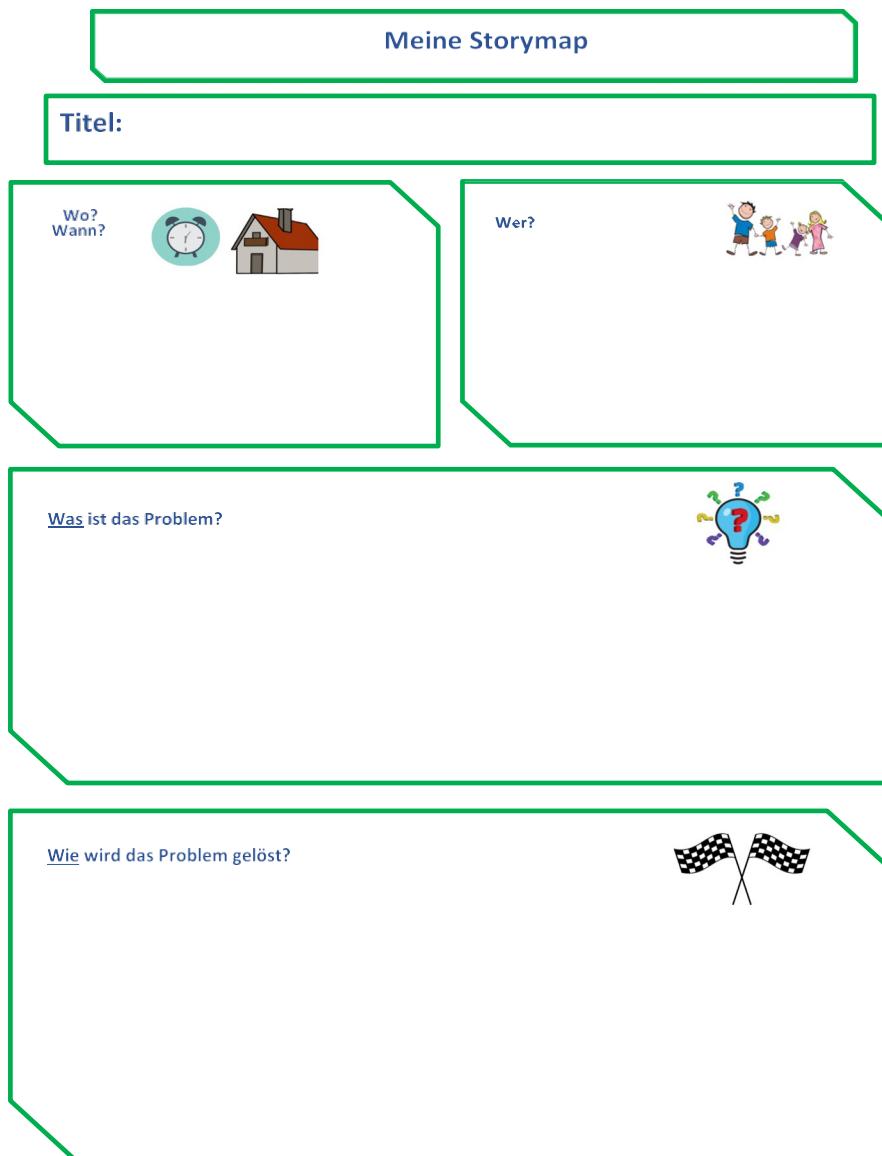
Note: My Story Map: Where? When? Who? What is the problem? How was the problem solved?

Intervention procedure

We evenly distributed the intervention sessions among the four interventionists so that two master's students were always present at each session. Furthermore, the class teachers were always present. We designed each support session to last 45 min. We developed the individual sessions and respective procedures in advance in small steps based on an intervention support plan. The SRSD model presented in the introduction served as a basic building block for the intervention structure and procedure. We ran through the six phases of the SRSD method during the course of the sessions to teach the students how to understand a text. We divided each session into different components and application phases. The students created a new story using a story map. Each session began with a welcome. The interventionists presented a direct instruction on the story map at the beginning of the intervention. During this phase, the interventionists familiarized the students with the story questions and the structure of the story map. Later, this initial phase also served to reactivate prior knowledge from the previous units. We then divided the 12 sessions into three categories for the application phase. We assigned the first two sessions to the I do it phase. We assigned Sessions 3–6 to the we do it phase, and the final sessions, beginning with Session 7, to the you do it phase. Common to all the phases was a final distribution of stickers or stamps for good cooperation and a farewell. As previously mentioned, the SRSD method aims to provide children with strategies they can apply independently. The support began with the I do it phase, which primarily focuses on developing background knowledge, modelling, and discussion. The interventionists initially presented the story map to the students in small steps using story cards. In the next step, the interventionists read a story aloud, using a beamer so that the children could read along and complete the story map as a group. The interventionists then completed the story map on the board as an example. The interventionists clearly communicated the individual steps, making the procedure transparent. In the transition to the we do it phase, the children became increasingly involved in the subsequent units and received guidance to complete the story maps together. The interventionists constantly repeated the story-map and story-card components to make the experience memorable. As the support units progressed, the interventionists ultimately guided the students toward the you do it phase, during which, after reading the story in plenary as a group, the children completed the story map with a partner in their My Reading Folders. Following the completion of the you do it phase, the interventionists compiled the results of the partner work and discussed any discrepancies.

We used positive reinforcement as a reward system to promote students' motivation throughout the intervention phase. We developed a reinforcement plan where the students

received a stamp or sticker after each session if they had worked well. The interventionists checked the story maps and awarded a sticker for a meaningful completion. Once they had earned four stickers, the children received a piece of candy from a reward box. In addition, the interventionists consistently delivered positive feedback during the sessions. The students also received "Reading Professional" certificates after the 12 sessions to mark the successful completion of the intervention.

Control condition

The control group attended regular lessons and did not receive explicit support in reading and writing. While the treatment group participated in the intervention, the control group had either a mathematics, art, or science lesson, depending on the day. We explicitly asked the teachers not to additionally foster reading and writing.

Social validity

It is necessary to evaluate the social validity (Luiselli & Reed, 2011) of intervention procedures to assess the acceptability of an intervention by those involved. We used rating scales for the assessment. In this study, we employed a five-point Likert scale and created two social validity forms: one for the participants and one for the teachers. The respondents answered questions with a positive valence by selecting an appropriate response: (1) do not agree at all, (2) disagree, (3) no opinion, (4) agree, and (5) fully agree. The social validity sheet listed the following items for the participants: (1) *The story map has helped me to understand texts better*, (2) *The story map has simplified story writing for me*, (3) I was happy to come to the intervention, (4) *I would like to do this more often*, (5) Peer tutoring has been fun and helpful, and (6) *The self-graphing and stickers were great and motivated me*. The sheet for the teachers listed the following items: (1) The children have learned a lot during the intervention, (2) *The children liked the intervention*, (3) I would implement the methods and the procedure in my classroom, (4) The intervention was appropriate in terms of level, (5) *I feel like the children's reading comprehension has improved*, and (6) *I feel like the children's text production has improved*.

Treatment fidelity

Treatment fidelity ensures that an intervention is implemented as designed; it includes adherence, quality, and exposure dimensions. Low fidelity can hinder interpretation of the results. An objective external individual should assess treatment fidelity. Treatment fidelity assesses the intervention's adherence (*Were all intervention elements adhered to?*), exposure or dosage (*Was the timescale for the intervention adhered to? Was the intervention delivered weekly as planned?*), and quality (*Were the intervention elements communicated as practiced?*).

The interventionists completed a checklist, and two external master's students (also present during the sessions) completed the treatment fidelity forms for each session. In addition, the interventionists received a detailed explanation of the treatment fidelity form. The two independent assessors rated the individual categories, compared them across the groups, and calculated the level of agreement. The interventionists who completed the forms merely served to provide information for the researchers. The external assessors rated the treatment fidelity as 100% (Sanetti et al., 2021). In addition, two assessors were present in the control groups to document the groups' activities and, above all, whether the class teacher explicitly taught reading and writing skills.

Data analysis

As the results of the pretests of the two groups exhibited differences with an effect size of > 0.05, statistical justification was necessary under WWC guidelines. Therefore, we conducted an analysis of covariance (ANCOVA) using the pretest as the covariate and the post and follow-up tests as dependent variables. We performed the ANCOVA for sentence reading, text reading, and text writing, controlling for pretest differences in the two groups. Before we conducted the ANCOVA, we checked the prerequisites for each dependent variable. We tested for homogeneity of regression to determine whether the within-group regression coefficients were equivalent.

<div align="center">Results</div>

Reading comprehension (sentence level)

The descriptive statistics for sentence reading demonstrate a definite improvement in reading comprehension on the sentence level in the story-map group compared to the control group. This finding appeared stable in the follow-up test. The homogeneity of the regression slopes was not violated regarding the dependent variable posttest, as the interaction terms were not statistically significant ($p > 0.05$). The main effect of the treatment condition proved to be statistically significant: $F(1,56) = 11.74$, $p < 0.01$, with an eta-square of 0.17. This result indicates that group membership had a significant effect on the posttest scores when controlling for the pretest differences.

Reading comprehension (text level)

The descriptive statistics for text reading reveal a demonstrable improvement in the story-map group compared to the control group. This finding also appeared stable in the follow-up test. The homogeneity of the regression slopes was not violated regarding the dependent variable follow-up, as the interaction terms were not statistically significant ($p > 0.05$). The main effect of the treatment condition proved to be statistically significant: $F(1,56) = 8.15$, $p <$

0.01, with an eta-square of 0.13. This result indicates that group membership had a significant effect on the follow-up scores when controlling for the pretest differences. Reading comprehension (text level) The descriptive statistics for text reading reveal a demonstrable improvement in the story-map group compared to the control group. This finding also appeared stable in the follow-up test. The homogeneity of the regression slopes was not violated regarding the dependent variable posttest, as the interaction terms were not statistically significant ($p > 0.05$). The main effect of the treatment condition proved to be statistically significant: $F(1,56) = 32.06$, $p < 0.001$, with an eta-square of 0.36. This result indicates that group membership had a significant effect on the posttest scores when controlling for pretest differences. The homogeneity of the regression slopes was not violated regarding the dependent variable follow-up, as the interaction terms were not statistically significant ($p > 0.05$). The main effect of the treatment condition proved to be statistically significant: $F(1,56) = 12.87$, $p < 0.001$, with an eta-square of 0.19. This finding indicates that group membership had a significant effect on the follow-up scores when controlling for pretest differences.

Text writing quality

Although the descriptive statistics reveal that the story-map group has improved, with higher results in the post- and follow-up tests than in the pretest, it cannot be clearly stated that the improvement is attributable to the intervention (even though pretest differences are controlled for). The homogeneity of the regression slopes was violated regarding the dependent variable posttest, as the interaction terms were statistically significant ($p < 0.05$). There is otherwise no corresponding countermeasure if this requirement is violated. Therefore, we can make no precise statements regarding the text writing quality, even though the main effect of the treatment condition proved to be statistically significant, $F(1,56) = 3.82$, $p < 0.05$, with an eta-square of 0.06, indicating a medium effect. The homogeneity of the regression slopes was not violated regarding the dependent variable follow-up, as the interaction terms were not statistically significant ($p < 0.05$). However, the main effect of the treatment condition proved not to be statistically significant: $F(1,56) = 1.31$, $p > 0.05$, with an etasquare of 0.02.

Social validity

Overall, the children and teachers felt very positive about the intervention, assigning the values 4 (agree) and 5 (fully agree) for each item. The children agreed that the story-map intervention had helped them better understand what they had read and write better stories. Furthermore, they enjoyed the activity and would like to participate in it more often. In addition, the children had the opportunity to leave comments. These clearly demonstrated that they were grateful and looked forward to the activity each time. The teachers also stated that the

intervention was very successful and that they would like to implement it in the classroom. Significantly, the teachers indicated that they felt that the students' overall reading comprehension and writing in the classroom improved.

Discussion

Main findings

The objective of the study was to assess the impact of a composite reading intervention on students facing reading challenges, including students with and without German as an L2. Evidence suggests that while reading interventions generally improve reading comprehension, strategy instruction is particularly advantageous for students with significant learning difficulties (Klingner et al., 2015). Comprehending text involves more than the ability to recognize words; it requires integrating and interpreting information (Oakhill et al., 2019). Additionally, Okkinga et al. (2018) have indicated that classroom-based interventions are especially effective for enhancing reading comprehension skills. The results of our study align with previous research (e.g., Cho et al., 2021; Okkinga et al., 2018; Sohn et al., 2023), as it demonstrates significant effects on reading indicated by a notable interaction effect with stable follow-up data. The observed improvement in comprehension highlights that reading proficiency is pivotal for academic achievement. This finding aligns with the objective of our study, which focuses on enhancing fundamental and advanced linguistic skills through tailored interventions. Sentence-level fluency is particularly crucial for L2 learners to grasp more complex text structures (.lva-rez-Ca.izo et al., 2015). Our findings concur with this observation, indicating that the intervention's effect was significant in terms of text comprehension.

The intervention appears to have had a more pronounced impact on text comprehension than on sentence comprehension. We can likely attribute this finding to the intervention's design, which enhances text-reading comprehension, providing the students with more extensive training in this area. Research conducted by Al-Otaiba et al. (2023) supports the efficacy of multicomponent interventions in bolstering reading comprehension at the upper primary level. Furthermore, a body of research, including a study by Graham et al. (2017), suggests a synergistic relationship between reading and writing skills, indicating that reading improvements can enhance writing abilities.

Therefore, we also measured writing competency to determine whether our intervention could simultaneously foster several important literacy components, which would be of great interest to schools and teachers. The study results offer a slight indication that the intervention could also have influenced text writing quality. Although we observed a significant main group effect when controlling for pretest differences, we can draw no clear conclusion because the

requirements for conducting an ANCOVA have been violated. Descriptively, a clear improvement is evident in the treatment group compared to the control group. However, these findings must be interpreted with caution and only possibly indicate that the intervention also promoted text writing quality. Nevertheless, our results contribute to the existing body of literature because research on writing skills, which we consider to be a crucial factor in educational attainment, is limited. The flexibility of our intervention is paramount, particularly for L2 learners within the K-12 educational spectrum. Research (e.g., Cho et al., 2021) has demonstrated that such interventions, when inclusive of strategic instruction, have significantly boosted reading proficiency among this demographic. The goal is to enhance sentence processing skills while also establishing a foundation for comprehending broader text structures. The results of this study corroborate the efficacy of this dual focus.

This study's findings suggest that the combined reading intervention had a positive impact on students with reading difficulties, with and without German as an L2, in terms of improving their reading, and possibly their writing skills, in a brief amount of time. These results support the findings of Sohn et al. (2023).

Furthermore, evaluating the combined reading intervention within the context of social validity yielded positive results. Social validity is an essential aspect to consider when assessing an intervention's overall effectiveness, as it reflects the extent to which the intervention can be viewed as relevant, acceptable, and beneficial. Therefore, the positive evaluation of the intervention suggests that it has the potential to be an effective and valuable tool for improving reading skills and outcomes for individuals (Luiselli & Reed, 2011). Integrating the intervention into classroom instruction is a critical factor bridging research with pedagogical practice. This approach directly addresses the challenge of translating intervention research findings into practical, actionable strategies that can aid educators and educational institutions. Notably, employing a multitiered reading system accommodating groups of 5–15 students has proven to be more effective than traditional small-group instruction. Furthermore, teachers reported observable advancements in student performance, not only within the target group but also among those not directly involved in the intervention, indicating the approach's potential overall efficacy. However, further data is required to substantiate these broader impacts within the general school context.

Study limitations and further research

This study has several limitations to consider when interpreting the results. Firstly, the sample size for each group was relatively small (30 participants in each group), which could limit the generalizability of the results to larger populations. However, finding large groups of

less proficient readers for an intervention study can be challenging. Nevertheless, future studies replicating this work using a larger student population would be beneficial for greater generalizability. A further limitation is that we did not compare the story map intervention with another reading intervention to determine which was more effective. Therefore, we can only confirm that the story-map intervention is effective in fostering reading comprehension and text production. Future research could include moderator variables, such as second language learners, individuals with special needs, gender, age, and behavior, to determine which components affect an intervention's effectiveness. Moreover, due to the connection between reading comprehension and reading fluency (De Jong et al., 2012), students could be screened for reading fluency in addition to reading comprehension to exclude those who might find the intervention too complex and difficult. It would also be interesting to explore whether the intervention's effectiveness also depends on the students' levels of reading fluency. The same idea applies to vocabulary since vocabulary significantly contributes to reading comprehension, especially in students learning an L2 (Lervåg & Aukrust, 2010). However, researchers should limit testing students who already experience a level of difficulty and failure in school. Finally, this study provides an indication of the influence of a reading intervention on writing ability but not vice versa. Thus, it would be interesting to follow Graham and Hebert (2011), who found that writing instruction facilitates reading. Another limitation is the limited reliability of the writing rubric; although our items achieved good values, the sample size was very small. The current lack of a standardized assessment of writing that can also compare scores with those of a normal sample should be considered when interpreting the results. Nevertheless, the results provide interesting indications of the potential influence of a reading intervention on writing, which are valuable in the context of schools and education. In this context, the correlation between the items should also be noted. Although we observed good to high correlations between the items within the conditions, item 5 fell outside the grid. This finding could be because the topic of emotions was a challenging rubric for many students.

Practical implications

Overall, the intervention is not difficult to implement in the classroom context, especially because it is adaptable to different students. One crucial aspect is to adapt the material; in addition to real-life stories, the texts must be written at the student's reading level to prevent frustration. If the teacher feels that the students are not keeping up, they can modify the individual phases of the support. For example, direct instruction for the story map should be longer if it is clear that students have not fully understood it. The intervention should only progress to more independent steps when knowledge of the story map has been consolidated. In the context

of the stories, it is also crucial that the students can clearly answer all the W-questions with the help of the story—otherwise, they could become confused. Assigning peers in the context of peer tutoring is equally important. Teachers should take care to ensure that the respective students have a pleasant relationship with one another so that they can work well together on a basis of trust. Since all the students in our study experienced reading difficulties, peer tutoring with similarly strong students was chosen. But there is also the possibility of pairing a less good reader with a better reader. However, the type of implementation highly depends on the students in the class, making the intervention valuable, especially in the context of inclusive education and the right to education for all.

Conclusion

This study provides valuable insights into the effectiveness of the story-map intervention on reading comprehension, and possibly on text writing quality, in students. Further research is needed to address the study limitations and confirm the generalizability of the findings. Nevertheless, this study provides a promising direction for future research on the use of the story map as an instructional tool to enhance students' literacy.

Appendix

Table 3

Correlation of scores within the condition

Group category	Item category	Statistical measure	Item category				
			Item 1	Item 2	Item 3	Item 4	Item 5
Control group							
	Item 1						
		Pearson	1	0.757**	0.702**	0.601*	0.067
		Sig		< 0.001	< 0.001	< 0.05	0.784
	Item 2						
		Pearson	0.757**	1	0.673**	0.477	0.347
		Sig	< 0.001		< 0.01	0.085	0.145
	Item 3						
		Pearson	0.702**	0.673**	1	0.640**	0.351
		Sig	< 0.001	< 0.01		< 0.01	0.141
	Item 4						
		Pearson	0.612*	0.477	0.640**	1	0.578**
		Sig	< 0.05	0.085	< 0.01		< 0.01
	Item 5						
		Pearson	0.067	0.347	0.351	0.578**	1
		Sig	0.784	0.145	0.141	< 0.01	
Story Map							
	Item 1						
		Pearson	1	0.802**	0.670**	0.600**	0.236
		Sig		< 0.001	< 0.001	< 0.01	0.303
	Item 2						
		Pearson	0.802**	1	0.606**	0.576**	0.087
		Sig	< 0.001		< 0.01	< 0.01	0.708
	Item 3						
		Pearson	0.670**	0.606**	1	0.548**	0.101
		Sig	< 0.001	< 0.01		< 0.01	0.662
	Item 4						
		Pearson	0.600**	0.576**	0.548*	1	0.448*
		Sig	< 0.01	< 0.01	< 0.01		< 0.05
	Item 5						
		Pearson	0.236	0.087	0.101	0.448*	1
		Sig	0.303	0.708	0.662	< 0.05	

**Significant at the 0.001 or 0.01 level, *significant at the 0.05 level

References

Al Otaiba, S., McMaster, K., Wanzek, J., & Zaru, M. W. (2023). What we know and need to know about literacy interventions for elementary students with reading difficulties and disabilities, including dyslexia. *Reading Research Quarterly, 58*(2), 313–332. https://doi.org/10.1002/rrq.458

Álvarez-Cañizo, M., Suárez-Coalla, P., & Cuetos, F. (2015). The role of reading fluency in children's text comprehension. *Frontiers in Psychology, 6*, 1810. https://doi.org/10.3389/fpsyg.2015.01810

Alves, K. D., Kennedy, M. J., Brown, T. S., & Solis, M. (2015). Story grammar instruction with third and fifth grade students with learning disabilities and other struggling readers. *Learning Disabilities: A Contemporary Journal, 13*(1), 73–93.

Bogaerds-Hazenberg, S. T. M., Evers-Vermeul, J., & van den Bergh, H. (2020). A meta-analysis on the effects of text structure instruction on reading comprehension in the upper elementary grades. *Reading Research Quarterly, 56*(3), 435–462. https://doi.org/10.1002/rrq.311

Bruggink, M., Swart, N., Van der Lee, A., & Segers, E. (2022). Putting PIRLS to use in class-rooms across the globe: Evidence-based contributions for teaching reading comprehension in a multi-lingual context (p. 110). Springer Nature. https://doi.org/10.1007/978-3-030-95266-2

Cain, K. (2022). Children's reading comprehension difficulties. In: M. J. Snowling, C. Hulme, & K. Nation (Eds.), *The Science of Reading*. Wiley. NY, USA.

Catts, H. W., Compton, D., Tomblin, J. B., & Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology, 104*(1), 166–181. https://doi.org/10. 1037/a0025323

Chen, J., Zhang, L. J., & Parr, J. M. (2021). Improving EFL students' text revision with the self-regulated strategy development (SRSD) model. *Metacognitive Learning, 17*, 1–21. https://doi.org/10. 1007/s11409-021-09280-w

Cho, Y., Kim, D., & Jeong, S. (2021). Evidence-based reading interventions for English language learners: A multilevel meta-analysis. *Heliyon, 7*(9), E07985. https://doi.org/10.1016/j.heliyon. 2021.e07985

Ciullo, S. P., & Reutebuch, C. (2013). Computer-based graphic organizers for students with LD: A systematic review of literature. *Learning Disabilities Research & Practice, 28*(4), 196–210. https://doi.org/10.1111/ldrp.12017

Clinton, V., Taylor, T., Bajpayee, S., Seipel, B., Carlson, S. E., & Davison, M. (2020).

Inferential comprehension differences between narrative and expository texts: A systematic review and meta-analysis. *Reading and Writing, 33*, 2223–2248. https://doi.org/10.1007/s11145-020-10044-2

Cure, G., Batu, E. S., & Gulboy, E. (2020). Effectiveness of the story-mapping strategy in students with disabilities: Meta-analysis of single-case experimental design studies. *Reading & Writing Quarterly, 37*(6), 513–534. https://doi.org/10.1080/10573569.2020.1858211

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

Gillespie Rouse, A., & Sandoval, A. (2018). Writing interventions for students with learning disabilities: Characteristics of recent research. *Learning Disabilities: A Multidisciplinary Journal, 23*(2), 1–17.

Grabe, W. (2016). L2 reading comprehension and development. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (vol. 3, pp. 299–311). Routledge, London UK. https://doi.org/10.4324/9781315716893

Graham, S., & Hebert, M. (2011). Writing-to-read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*(4), 710–744.

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., Barkel, A., Kavanaugh, C., & Talukdar, J. (2017). Reading for writing: A meta-analysis of the impact of reading interventions on writing. *Review of Educational Research, 88*(2), 243–284. https://doi.org/10.3102/00346 54317746927

Gunderson, L., & D'Silva, R. A. (2016). Second language literacy Kindergarten to grade 12. In: E. Hinkel (Ed.), *Handbook of research in second language teaching and learning*, Routledge, London, UK Han, J., & Hiver, P. (2018). Genre-based L2 writing instruction and writing-specific psychological factors: The dynamics of change. *Journal of Second Language Writing, 40*, 44–59. https://doi.org/10.1016/j.jslw.2018.03.001

Harris, K. R., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation* (2nd ed.). MA, USA: Brookline Books.

Hausman, N. L., Javed, N., Bednar, M. K., Guell, M., Schaller, E., Nevill, R. E., & Kahng, S. W. (2022). Interobserver agreement: A preliminary investigation into how much is enough? *Journal of Applied Behavior Analysis, 55*(2), 357–368. https://doi.org/10.1002/jaba.811

IDA Editorial Contributors. (2020). *Dyslexia basics*. International Dyslexia Association.

Klingner, J. K., Vaughn, S., & Boardman, A. (2015). *Teaching reading comprehension to students with learning difficulties, 2/E: What works for special-needs learners* (2nd ed.). NY, USA: Guilford Publications.

Lenhard, W., Lenhard, A., & Schneider, W. (2020). *ELFE II. Ein Leseverständnistest für Erst- bis Siebt- klässler – Version II [ELFE II. A reading comprehension test for first to seventh graders - version II]* (4th ed.) Hogrefe.

Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry, 51*(5), 612–620. https://doi.org/10.1111/j.1469-7610.2009.02185.x

Luiselli, J. K., & Reed D. D. (2011). Social validity. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of child behavior and development* (p. 1406). Springer. https://doi.org/10.1007/978-0-387-79061-9_3168

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.

Marzec-Stawiarska, M. (2016). The influence of summary writing on the development of reading skills in a foreign language. *System, 59*, 90–99. https://doi.org/10.1016/j.system.2016.04.006

McBreen, M., & Savage, R. (2020). The impact of motivational reading instruction on the reading achievement and motivation of students: A systematic review and meta-analysis. *Educational Psychology Review, 33*(3), 1125–1163. https://doi.org/10.1007/s10648-020-09584-4

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.

McLeskey, J., Barringer, M.-D., Billingsley, B., Brownell, M., Jackson, D., Kennedy, M., & Ziegler, D. (2017). *High-leverage practices in special education*. Council for Exceptional Children & CEEDAR Center.

Oakhill, J., Cain, K., & Elbro, C. (2019). Reading comprehension and reading comprehension difficulties. In D. A. Kilpatrick, M. R. Joshi, & R. K. Wagner (Eds.), *Reading development and difficulties: Bridging the gap between research and practice* (pp. 83–115). Springer. https://doi.org/10.1007/ 978-3-030-26550-2_5

Okkinga, M., van Steensel, R., van Gelderen, A. J. S., van Schooten, E., Sleegers, P. J. C., & Arends, L. R. (2018). Effectiveness of reading-strategy interventions in whole classrooms: A meta-analysis. *Educational Psychology Review, 30*(4), 1215–1239. https://doi.org/10.1007/s10648-018-9445-7

Onachukwu, I., Boon, R. T., Fore, C., & Bender, W. N. (2007). Use of story-mapping procedures in middle school language arts instruction to improve the comprehension skills of students with learning disabilities. *Insights in Learning Disabilities, 4,* 27–47.

Organisation for Economic Cooperation and Development. (2019). *Programme for international students assessment (PISA) results from PISA 2018.* https://www.oecd.org/pisa/publications/PISA2018_CN_ DEU.pdf

Rezaei, A. R., & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing, 15*(1), 18–39. https://doi.org/10.1016/j.asw.2010.01.003

Sanders, S. (2020). Using the self-regulated strategy development framework to teach reading comprehension strategies to elementary students with disabilities. *Education and Treatment of Children, 43*(1), 57–70. https://doi.org/10.1007/s43494-020-00009-z

Sanders, S., Losinski, M., Parks Ennis, R., White, W., Teagarden, J., & Lane, J. (2019). A meta-analysis of self-regulated strategy development reading interventions to improve the reading comprehension of students with disabilities. *Reading & Writing Quarterly, 35*(4), 339–353. https://doi.org/10.1080/ 10573569.2018.1545616

Sanetti, L. M. H., Cook, B. G., & Cook, L. (2021). Treatment fidelity: What it is and why it matters. *Learning Disabilities Research & Practice, 36*(1), 5–11. https://doi.org/10.1111/ldrp.12238

Sencibaugh, J. M. (2007). Meta-analysis of reading comprehension interventions for students with learning disabilities: Strategies and implications. *Reading Improvement, 44*(1), 6–22.

Shin, J. (2020). A meta-analysis of the relationship between working memory and second language reading comprehension: Does task type matter? *Applied Psycholinguistics, 41*(4), 873–900. https://doi. org/10.1017/S0142716420000272

Sohn, H., Acosta, K. R., Brownell, M. T., Gage, N. A., Tompson, E., & Pudvah, C. (2023). A meta-analysis of interventions to improve reading comprehension outcomes for adolescents with reading difficulties. *Learning Disabilities Research and Practice, 38*(2), 85–103. https://doi.org/10.1111/ ldrp.12307

Stagliano, C., & Boon, R. (2009). The effects of a story-mapping procedure to improve the comprehension skills of expository text passages for elementary students with learning

disabilities. *Learning Disabilities: A Contemporary Journal, 7*(2), 35–58.

Sun, Y., Wang, J., Dong, Y., Zheng, H., Yang, J., Zhao, Y., & Dong, W. (2021). The relationship between reading strategy and reading comprehension: A meta-analysis. *Frontiers in Psychology, 12*, 635289. https://doi.org/10.3389/fpsyg.2021.635289

Tang, S., Irby, B. J., Tong, F., & Lara-Alecio, R. (2021). The effects of cooperative, collaborative, and peer-tutoring strategies on English learners' reading and speaking proficiencies in an English-medium context: A research synthesis. *SAGE Open, 11*(4), 21582440211060824. https://doi.org/10. 1177/21582440211060823

Troia, G. A., & Graham, S. (2002). The effectiveness of a highly explicit, teacher-directed strategy instruction routine: Changing the writing performance of students with learning disabilities. *Journal of Learning Disabilities, 35*(4), 290–305. https://doi.org/10.1177/00222194020350040101

Zhang, J., & Wang, Q. (2023). Reading and writing difficulties in bilingual learners. *Annals of Dyslexia, 73*(1), 1–5. https://doi.org/10.1007/s11881-023-00282-8

Appendix D Article 4 *(peer-reviewed)*

Disclaimer for Dissertation Version:

This is the accepted manuscript of the article listed above. The final published version is available at the DOI provided.

Abstract

The number of students learning German as a second language (L2) is steadily increasing. Unfortunately, studies reveal that less-proficient school performance affects a larger proportion of these students and additional behavioral problems can create even greater learning barriers. In order to master a language, the focus is not only on vocabulary, but also on reading, and studies show that multi-component intervention in reading and L2 acquisition is particularly promising. Therefore, this multiple baseline study focuses on a multi-component storytelling intervention on vocabulary, reading, and letter sound fluency of low-achieving first graders with German as L2 with and without behavioral problems ($N = 7$). The intervention was implemented 3 times a week over a 6-week period. Results show significant large to very large effects on vocabulary and moderate to large effects on letter sound fluency and reading, providing indication for the positive impact of storytelling on multiple aspects simultaneously for the focused sample.

Keywords: Storytelling, Vocabulary, Reading and Letter Sound Fluency, German Second Language, Behavior Problems

Introduction

German as a Second Language

Education is largely dependent on language and in the German education system, the understanding and speaking of German at native language level is assumed (Becker-Mrotzek et al., 2012). According to the Federal Statistical Office, about 11% of the students at educational institutions have a migration background (Federal Statistical Office, 2020) and learning

German as second language (L2; Aschenbrenner et al., 2016). The Programme for International Student Assessment (PISA) studies shows that students with a migration background perform significantly worse at school than students who learn German as their mother tongue [Organisation for Economic Cooperation and Development (OECD), 2019], and it has been shown that a large proportion of fourth graders do not or only partly speak German at home (Hußmann et al., 2017). German L2 students struggle in schools leading to a challenge for the teachers in designing appropriate lessons (Becker-Mrotzek et al., 2012) and a challenge for the students themselves with respect to educational opportunities.

Hurdles for Second Language Learners

To be proficient in language, various skills within language acquisition, such as phonetics and literacy, are needed (Aschenbrenner et al., 2016). But especially vocabulary learning is immensely important (Schmitt et al., 2011), and it is shown that particularly students with L2 experience severe failure in this area (Webb and Chang, 2012). In addition, letter sound fluency (LSF) is essential for language communication and acquiring the sound of individual letters presents a particular hurdle (Kim and Piper, 2019), and students who struggle with LSF are more likely to have difficulty in their reading skills later on as well (Piasta and Wagner, 2010). A reason for this might be that children fail to read because their overall L2 competence is not yet sufficient to read adequately (Wallace, 2014).

Also, it is widely known that a certain struggle in language development, as vocabulary, expression, and reading, can be associated with problems in behavior (Peterson et al., 2013; Jansen et al., 2020). It has been reported that young children with language difficulties might develop problem behavior (Henrichs et al., 2013) which can get worse over time (Curtis et al., 2018). More specifically, deficits in language are connected to deficits in attention processing (Peterson et al., 2013) which can be linked to learning-related behavior (LRB). LRB, according to McClelland et al. (2006), includes abilities like staying focused, organizing school material, and working on one's own. A meta-analysis by Chow and Wehby (2018) revealed a negative relationship between language deficits and problem behavior independent of age and time.

Important Language Components

Vocabulary is fundamental but challenging in a L2 language and influences all stages of acquisition (Ender, 2016). Vocabulary can be differentiated between expressive and receptive. Receptive vocabulary is words which can be recognized but not actively spoken, whereas expressive vocabulary can be directly used (Schmitt, 2014). Significant correlations have been shown to exist between expressive vocabulary and reading ability in children from primary school (Wise et al., 2007). In general, it has been found that L2 vocabulary knowledge is linked

to L2 reading comprehension (Lervåg and Aukrust, 2010). As in an L1, the automation of lower hierarchy processes, such as word recognition, is fundamental for comprehension (Kramer and McLean, 2019). The Dual-Route Model (DRM; Coltheart, 2005) describes two routes, the lexical and the non-lexical route, to show how readers read aloud. The lexical route refers to the mental lexicon where words can be automatically stored and retrieved [more important for irregular words: e.g., *"hoch" (high)* than for regular words: e.g., *"Sand" (sand)*]. The non-lexical route goes through the grapheme-phoneme correspondence (e.g., important for non-words like "*brelo*" or "*blustof*"). In terms of direct word recognition, the direct route is important, where sight words can be retrieved. Sight words are words that can be retrieved within 1 s of occurrence (Ehri, 2005). In addition to memorizing familiar words, letter sound knowledge (LSK) also plays an important role in the non-lexical route of DRM because it enables readers to decode unfamiliar words (Ehri, 2002). Both approaches should be possible for a reader to build up adequate reading competence in a language. Clemens et al. (2017) found that LSF, a sub-component, was predictive of subsequent reading fluency with respect to kindergarten children. Through a mediation analysis of results from a large-scale intervention study ($N = 152$), Hulme et al. (2012) showed that problems in LSK and phoneme awareness can cause difficulties in later word-reading-proficiency in 5-year-old children.

Fostering Second Language Acquisition

In order to counteract hurdles in second language acquisition and to offer L2 students an opportunity to acquire an L2 adequately, it is necessary to provide effective support. The dual-coding theory (DCT; Paivio, 2008) states that there is a verbal way and a non-verbal way (i.e., pictures) to store information underlining the importance of presenting new input verbally and non-verbally in a language, especially for L2 students (Huang et al., 2019). The verbal way is related to linguistic information (e.g., sound) and the non-verbal system is linked to visual information (e.g., pictures; Paivio, 2007). According to Reed (2010) using both systems, maximizes the likelihood that information will be stored adequately. Another way to train new content is either through explicit (intentional) training or implicit (incidental) training (Jin and Webb, 2020) – or a combination of both (Choo et al., 2012). Intentional learning means that the learner is aware of learning something, and incidental learning means that the learner learns something like a by-product without being aware of it (Webb and Nation, 2017). In the case of incidental learning, it has been said that words are easier to acquire through repeated occurrence in context (Webb and Nation, 2017). Marulis and Neuman (2010) conducted a meta-analysis about the impact of vocabulary interventions on the language development of pre-K and kindergarten children and found an overall effect size of $g = 0.88$ of vocabulary training on word

learning. Moreover, it was found that a combination of implicit training and explicit training lead to a higher effect size ($g = 1.21$) than explicit ($g = 1.11$) and implicit ($g = 0.62$) in isolation. Hulme et al. (2012) found that teaching LSK and phoneme awareness explicitly in a reading and phonology intervention lead to an improvement of these two abilities.

It has been shown that multi-component supports, including among other, phonics, vocabulary, and fluency increases the probability of becoming a good reader (Foorman and Torgesen, 2001). A recently published literature review by Donegan and Wanzek (2021) showed that multi-component reading interventions for elementary school with the highest effects incorporate instruction in decoding on the word level and in summary multi-component interventions are promising with regard to improve foundational reading skills and reading comprehension.

Storytelling

Listening to stories has been known to influence language development on different areas positively in children (Isbell et al., 2004). Storytelling is a procedure where a teller tells a story in an authentic environment using gestures, vocalization, and images to convey a certain message to the audience who are incorporated in the storytelling procedure (Mello, 2001). Storytelling has the ability to engage learners personally (Brewster et al., 2002), motivate learners, and spark interest in the subject matter (Wright, 2013). Using storytelling does have positive impacts on child's oral and written language development (Fien et al., 2011; Baker et al., 2013) and through the procedure of storytelling facts as well as vocabulary can be memorized better (Wajnryb, 2003). Lenhart et al. (2018) focused on the impact of story listening on vocabulary acquisition and found that vocabulary was acquired incidentally without any word explanation with a moderate effect ($d = 0.37$) which was in turn not stable over time (age 3–6) concluding that using only incidental vocabulary training might not be sufficient enough. A meta-analysis by Mello (2001) indicates that using storytelling led to gains in vocabulary, fluency, and writing skills, among other variables. Suggate et al. (2013) examined storytelling in second and fourth grade German readers and revealed that more freely storytelling has more benefits than simply reading the story.

Read aloud has been shown to be effective for vocabulary, comprehension, and narrative language in first graders (Baker et al., 2020) and for phonological awareness (Swanson et al., 2011). Since storytelling belongs rather to the implicit method, adding flashcards to storytelling in order to teach components explicitly would be, according to Marulis and Neuman (2010), a further boost in effectiveness. Two additional studies by Barwasser et al. (2020) and Knaak et al. (2021) investigated a combined storytelling intervention consisting of implicit and

explicit components on vocabulary acquisition in English language learning of students with and without learning disabilities showing that this combination is effective in the context of vocabulary acquisition. Barwasser et al. (2021) went a step further and examined the combined storytelling method in German second language learners from primary school on vocabulary and reading with overall positive effects.

Motivation and Self-Graphing

In second and foreign language acquisition, the ability to increase competence in a language often depends on how motivated a learner is (Ghenghesh, 2010). Adding motivational components to an intervention can be specifically successful (Bowman-Perrott et al., 2013; Leko, 2016). It has been shown that incorporating self-monitoring procedures, such as self-graphing, the visualization of a student's own progress showing earlier scores and current scores (Stotz et al., 2008; Guzman et al., 2018; McKenna and Bettini, 2018), reading achievement can be improved for students with disabilities (Laurice and Eveleigh, 2011) and on task behavior as well as general academic productivity (DiGangi et al., 1991). Self-graphing can be realized by providing students with a graph overview where they can enter their scores after each measurement point in order to follow their own learning progress step by step. A meta-analysis by Guzman et al. (2018) revealed large effects of self-monitoring procedures on reading performance in K-12 students ($TauU = 0.79$, $p < 0.001$).

Research Questions

Based on the knowledge that there is a large number of low-performing German as a L2 language students in Germany, with both behavioral problems and motivation playing a significant role, a multi-component storytelling intervention was designed to simultaneously address three important components in language learning: vocabulary, LSF, and sight word reading, and to investigate its effects on German L2 students with and without behavioral problems. In addition, we have implemented a social validity questionnaire in order to figure out the acceptance of the intervention rated by the participants. Assessing social validity is a necessity to crystallize the acceptance and usefulness of interventions (e.g., Briesch et al., 2013). Accordingly, the four research questions are as follows:

1. Does a multi-component storytelling intervention lead to an increase in expressive vocabulary in German L2 students with and without behavior problems?

2. Does a multi-component storytelling intervention lead to an increase in LSF in German L2 students with and without behavior problems?

3. Does a multi-component storytelling intervention lead to an increase in sight word reading in German L2 students with and without behavior problems?

4. How was the intervention evaluated by the participants in terms of social validity?

Materials and Methods

Participants and Setting

The present study was conducted at an inclusive elementary school in a large city in North Rhine-Westphalia, Germany, targeting grade 1. To participate in the study, teachers of the respective classes were to identify all students who met the criterion "German as a second language" ($N = 10$). In addition, appropriate parental consent to participate in the study had to be obtained. The teachers received a teacher questionnaire to obtain relevant information on the proposed students regarding socio-demographic characteristics (see Table 1).

Table 1

Characteristics of the participants

	Lio	**Kim**	**Tila**	**Nele**	**Niek**	**Abden**	**Elif**
Age	6;3	6;5	7;1	6;6	6;3	7;1	6;2
Grade	1	1	1	1	1	1	1
Gender	Male	Female	Female	Female	Male	Male	Female
L1	Polish	Polish	Turkish	Chinese	Italian	Turkish	Turkish
LRB	12	4	13	3	14	10	6
Reading W (PR)	<7	<7	<2	9–13	7–8	16–17	5–11
Reading P (PR)	<2	<2	<4	24	6–8	19–23	8–10
BAKO (PR)							
Subtest PS	2	2	2	2	48	2	21
Subtest VS	42	6	19	53	6	6	61
Subtest RD	3	3	3	21	3	3	34
Subtest PI	8	8	8	74	8	8	21
Subtest SC	28	15	42	71	1	28	28
Subtest VD	57	7	23	57	23	23	57
Subtest WR	9	9	9	35	23	9	35
Total	7	0	2	39	2	1	31
Vocab (PR)	12	5	15	26	21	27	16

PR, percentile; W, words; P, pseudowords; LRB, learning-related behavior (cutoff 10); L1, native language; PS, pseudoword segmentation; VS, vowel substitution; RD, residual word determination; PI, phoneme interchange; SC, sound categorization; VD, vowel length determination; WR, word reversal; and Vocab, German vocabulary test.

German Vocabulary Test

A vocabulary test (WS/ZF-R; Weiß, 2007) in the form of a group screening was used first to assess the students' verbal language skills. The WS/ZF-R measures colloquial

vocabulary beyond the basic vocabulary of the German language and is used to determine the developmental level of verbal skills of students. The test sheet contains 30 multiple-choice items with five alternative answers each. Each task consists of a key word being given first. Subsequently, the respondents have to select the word from the five alternative answers that has a similar meaning as the given keyword. The reliability of the WS/ZF-R was assessed using the split-half method ($N = 618$), where values ranged from rtt = 0.79 to rtt = 0.90 with a mean value of rtt = 0.87. For the correlation with German grades ($N = 689$), the value was $r = 0.48$ (Weiß, 2007). The results are shown in percentiles (PR) and a PR under 15 means underdeveloped. For example, a percentile of 15 means 15 percent of the subjects in the norm sample scored the same or fewer points. The participant with a PR of 15 therefore belongs to the 15 percent of the weakest in his age group.

SLRT II

The Salzburg Reading and Spelling Test (SLRT II; Moll and Landerl, 2010) was used to assess reading ability at the word and pseudoword level. These two subtests each consisted of a one-minute reading fluency test by reading given words and pseudowords. The total time required is time-efficient at approximately five minutes. The parallel test reliability ranges from 0.90 to 0.98 and correlations with other German reading tests range from 0.69 to 0.92. All participants who had a percentile below 15 were selected for the study.

Test for Phonological Awareness (BAKO 1–4)

A test for phonological awareness for grades 1–4 was additionally used (BAKO 1–4; Stock et al., 2017). There are a total of 174 tasks divided into seven subtests: (1) pseudoword segmentation, (2) vowel substitution, (3) residual word determination, (4) phoneme interchange, (5) sound categorization, (6) vowel length determination, and (7) word reversal. The time required to complete the test is approximately 30 min. Norms are available for each grade level ($N = 876$) and reliability shows that internal consistently varies by grade level (between $\alpha = 0.90$ and $\alpha = 0.92$, split-half reliability between $r = 0.90$ and $r = 0.94$). Criterion-related validity with reading or spelling performance measured by standardized tests varies by grade level between $r = 0.42$ and $r = 0.68$ (Stock et al., 2017). Results are again shown in PR.

Integrated Teacher Report Form

The integrated teacher report form (ITRF; Volpe et al., 2018) represents a multilevel screening procedure used to identify student behavior difficulties. In the present study, the ITRF-G short version is applied, which is the German translation of the English version. In the research conducted, the screening is conducted by the classroom teachers as they are in the best position to assess the students' behavior. The teachers assess specific behaviors of the students

on an assessment sheet, and the items are created based on the factors "learning-related behavior" and "oppositional/disruptive behavior." The ITRF-G is administered in a short version with 16 items, whereas the original version includes 47 items. The conducted short version has been positively evaluated and shows high internal consistency and sufficient test-retest reliability in terms of reliability and high external validity for all scales in terms of validity. The cutoff value for learning-related behavior is 10 showing problems in this area (Volpe et al., 2018). Word Pretesting To crystallize the final training words and to ensure that the words were not stored in either the expressive vocabulary or the mental lexicon for reading, words were auditioned prior to the study. Once for expressive vocabulary and once for reading. The pool of words (N = 143) came from the Metacom symbols (Kitzinger, 2020) and care was taken to ensure that words were taken which the children could use well in everyday life. These words were queried both expressively and in reading. For the reading test (day 1), the 143 words were integrated into a powerpoint presentation so that one word was on each slide individually. The slides were scrolled in 1-s intervals, since according to Ehri (2005), a word is considered a sight word if it can be read within 1 s of its occurrence. Here, all words that could not be read were marked. After a few days (day 2), the expressive test was performed with the exact same words. Here, the children were not shown the 143 words, but pictures matching the words. Here, too, there was a picture on a slide – there was no time limit. Now, for each picture, the children were asked what the word was called. All non-conscious words were marked and compared with the reading words. A total of 40 word-overlaps resulted for unknown expressive words and words not read correctly. The 40 training words in reading were the same as in vocabulary for the intervention and measurements later on. Thus, the children could neither read these words nor express them actively. The 40 training words, which were selected together with the teachers, had a mid-frequency of $M = 10.5$, meaning that the words appear 10.5 times per million words in a corpus (Brysbaert et al., 2018). To estimate the frequency, we used the childLex database (Schroeder et al., 2015).

The students ($N = 10$) are divided into three groups. Group 1 had three children, group 2 had three children, and group 3 had four children. All participants learned German with the entry of kindergarten at age 3–3; 5. According to COVID-19 rules, groups were not allowed to be mixed across classrooms. Each group has a different baseline time and thus starts the intervention with a time delay. Three children are dropped from the data because they have too much missing data due to COVID-19 quarantine regulations. As a result, the finale sample for this paper is $N = 7$.

Design

The present research utilized a multiple baseline design across participants to examine the effects of the intervention. A single case analysis is often understood to be a study of one individual. However, a multiple baseline design embeds subcases within an overall case. The introduction of the intervention is temporally staggered across the subjects. The goal of implementing a multiple baseline design is to substantiate a cause-effects relationship by demonstrating that changes in the dependent variable only occur when the treatment is given (Lane et al., 2017). First, a baseline of varying length is performed with 5–7 sessions. After each of these sessions, the dependent variables were collected. After completion of the baseline phase, the intervention starts in the following sessions. Data were also collected after each intervention session (e.g., baseline 1 – measurements; baseline 2 – measurements – … intervention 1 – measurements; intervention 2 – measurements; and intervention 3 – measurements). Each group was randomly assigned to a specific baseline length resulting in group 1 = 5 baselines, group 2 = 6 baselines, and group 3 = 7 baselines. The baseline and intervention sessions took place 3 times a week for 25 min, after which the children were measured individually for each of the three dependent variables. The entire period spanned 6 weeks and 1 week of diagnostic testing. Due to a previous school closure because of COVID-19, the study started later and comes to 18 measurement time points of originally planned 24. Two master's students for special needs education functioned as test leaders and interventionists. Both supported each group together.

Dependent Variables and Measurement

In total, there are three dependent variables: expressive vocabulary, sight word reading, and LSF. The 40 training words were used for expressive vocabulary and reading. For LSF, all letters from the German alphabet were measured.

1. **Expressive vocabulary**: The 40 training words were packed into a powerpoint presentation in the form of pictures, with one picture per slide. For each picture, the child was asked if he knew the name of the word. The total number of correctly conscious words expressive was transferred to an excel table per measurement point.

2. **LSF:** All letters of the German alphabet were mixed and written on two 8.3 × 11.7-inch sheets, so that a total of 104 letters could be seen. The child was now asked to pronounce as many sounds as possible correctly within 1 min. A timer was set to 1 min and the two test leaders listened attentively. The total number of correctly pronounced sounds was also entered in the excel table for each measurement point.

3. **Reading:** The 40 training words written were embedded in a powerpoint. Here, one word per slide was written down. The slides were separated by hashtags and were laid out in 1-s intervals (see Ehri, 2005). Again, the total number of correctly read words was recorded in an excel table per measurement time point.

Intervention Material

For the direct instruction of the words and the sounds, a phonetic table and 8.3 × 11.7-inch flashcards with the letters on them and 8.3 × 11.7-inch flashcards with the picture and the matching word were used. For the storytelling intervention, short stories were required for each session. Before the study started, the master students talked to the children about their interests in order to determine the focus of the stories. In total, there was one full story with sub-chapters per session. The stories (example Figure 1) were self-written with somewhat the same length and formatting. Additionally, care was taken to ensure that all words occurring were not too difficult. The training words were always embedded and from the pool of 40 words always five were taken into one story which appeared twice on one story. The words were randomly assigned to the stories, making sure that in the end the words occurred in equal proportions. The training words in the story were always highlighted in blue, while the rest of the font was black.

Regarding the motivational system, there were three self-graphing sheets for the children corresponding to the three dependent variables. Each sheet consisted of several rows one below the other, which were supposed to represent the sessions (example Figure 2). The rows consisted of small boxes that were supposed to represent the number of words/sounds correctly known where the participants were asked to color the amount of correct known words/sounds after each measurement point.

Figure 1

Example part of a story

The secrets of the underwater world. The first day of school. Text: Alvin has a stomach-ache because he is excited. Tomorrow he is supposed to go to school for the first time. Of course, he is looking forward to it. Finally, he belongs to the big kids and is allowed to learn something, but he is also a little worried.

Figure 2

Example of self-graphing sheet

Wörtermeister = word master.

Procedures

Baseline

The baseline (A phase) is used to record the actual state in a multiple baseline design. Before the storytelling intervention starts, all three groups go through a baseline phase of different lengths for the groups. The baseline activities must not have anything to do with reading, vocabulary, or LSF, so that the dependent variables are not already promoted in the baseline. Thus, during baseline condition, games, puzzles, and math problems are solved together in 25 min. These are simple tasks that do not explicitly promote vocabulary, reading, or the LSF. Afterward, the three dependent variables are measured for each child.

Storytelling

After the baseline (A) phases are all completed, the groups begin the intervention phase (B phase). The group constellations remain unchanged. Storytelling can be divided into two stages. In the first stage (10 min), the kids sit in a semi-circle around the interventionist who is firstly introducing the words to be learned directly to the participants. Both, the words and letters of the last story (despite session one), are repeated, and the words and letters of the current story are introduced through flashcards and a phonics table. In order not to overtax the children, only 10 of the 40 words are directly instructed per session. The interventionist holds up the flashcard with the word and the picture, covers the written word, and asks the children,

based on the picture, whether they know what it means. Then, they talk about the word. Next, the interventionist uncovers the written word and asks the children if anyone can read the word aloud. Subsequently, everyone reads together and then, the interventionist reads the word again. After that, the interventionist lifts up the phonics picture. For each intervention session, 10 sounds were randomly selected to be trained. Using the phonics picture and the words, the interventionist asks, for example, for an "L": "Who knows how to pronounce that?" "And can you find the sound in one of our words?" The procedure lasts 10 min.

The second stage (15 min) involves the process of storytelling. The stories were learned by heart by the interventionists and the text serves the children to follow the story and see the marked training words. Each story is told out loud to the students and if a training word is appearing in the story, the story is paused and the word, as well as one sound, is discussed using the appropriate flashcards (a word with a matching picture). After the storytelling, the three measurements are carried out with each child individually and feedback on the learning process follows on the self-graphing sheets. Each time after the measurement, each child enters the number of correct known items in two separate self-graphing sheets for the amount of correctly read words and correctly known word expressively.

Treatment Fidelity

In order to record treatment fidelity in the present study, the experimenters were first provided with a detailed script with steps to be followed. Additionally, the implementers were given a checklist to complete at the end of each intervention session without being aware of the intention of the sheet. This was used to reflect on compliance with what was outlined in the script. The checklist is divided into six sections: environment/ external circumstances, planning, materials, procedure of support, diagnostics/feedback, and handling student behavior during support using three response options ("yes"; "no"; "not applicable"). In addition, a free field was available to the investigators for comments on special features in the context of the support. The inter-rater reliability is 100%.

Social Validity

To measure the acceptance of the intervention by the students, a questionnaire was designed within the framework of social validity, which was handed out to the students at the end of the support. The interventionists were not present in order to avoid biased results and to obtain an honest opinion from the students. The questionnaire contains nine items which should be rated on a scale from 0 (= completely not agree) to 4 (= completely agree). The items were as follows: (1) *Storytelling helped me to be able to read words correctly*; (2) *Storytelling helped me learn words and their meanings*; (3) *Storytelling helped me to pronounce sounds correctly*;

(4) *I understood well the meaning of the promotion*; (5) *I have learned a lot during storytelling*; (6) *I gladly came to the intervention sessions*; (7) *The self-graphing sheets were fun*; (8) *The stories were great*; and (9) *I would like to do more with stories in school*.

Data Analysis

The entire data analysis was done using the statistics program "R" and the Scan Package for multiple baseline design analysis in order to estimate the intervention (B phase) effects compared to the baseline (A phase). The graphs (Figures 3–5) for each dependent variable serve for visual analysis. In addition, mean and median values of the two phases as well as the maximum values in phase A and phase B were determined and mean baseline difference (MBDi). MBDi is a non-parametrical method which measures increase of a certain output from baseline (O'Brien and Repp, 1990). Further, overlap measures were used including the non-overlap of all pairs (NAP, Parker et al., 2011a), the percentage exceeding the median (PEM, Ma, 2006), the percentage of all non-overlapping data (PAND; Parker et al., 2007), and finally, the Tau-U additionally considering an A phase trend using the formula: A vs. B + TrendB − _TrendA. TauU measures data non-overlap between phase A and phase B (Parker et al., 2011b).

The single-case reporting guidelines by Tate et al. (2016) suggest the use of inferential statistics to directly test for treatment effects. Even though there is still no universal gold standard for analyzing data from respective experiments, hierarchical piecewise regression modeling has become the most common tool for investigating the null hypothesis (Raudenbush and Bryk, 2002; Manolov et al., 2010; Waddell et al., 2011). In this approach, the data points during baseline of one individual are used to calculate a regression line and estimate the progression of the data during the intervention. Changes in level and/or slope across phases can then be tested for statistical significance (level 1 analysis). Subsequently, data over several individuals can be accumulated to examine causal elements behind treatment effectiveness (level 2 analysis). When regression modeling is used in group studies, each data point stems from a different individual. However, if this approach is applied in single case level 1 research, the data points stem from one and the same person. One of the basic requirements for using parametric statistics (like regression analysis) is the independency of the distributed errors. There is no logical reason to assume that errors of different individuals are statistically associated. In contrast, the danger of autocorrelation in single case research is ever present. For example, it is anything but unlikely that errors in observations that are close together in time are more similar than those that are more distant. The degree to which they correlate corresponds with the risk of incorrectly rejecting a true null hypothesis. To reduce the likelihood of mistakenly dismissing the

absence of a given effect, we used a statistical package for R called SCAN (Wilbert, 2021) that controls for autocorrelation in single case data.

Figure 3

Amount of known expressive vocabulary
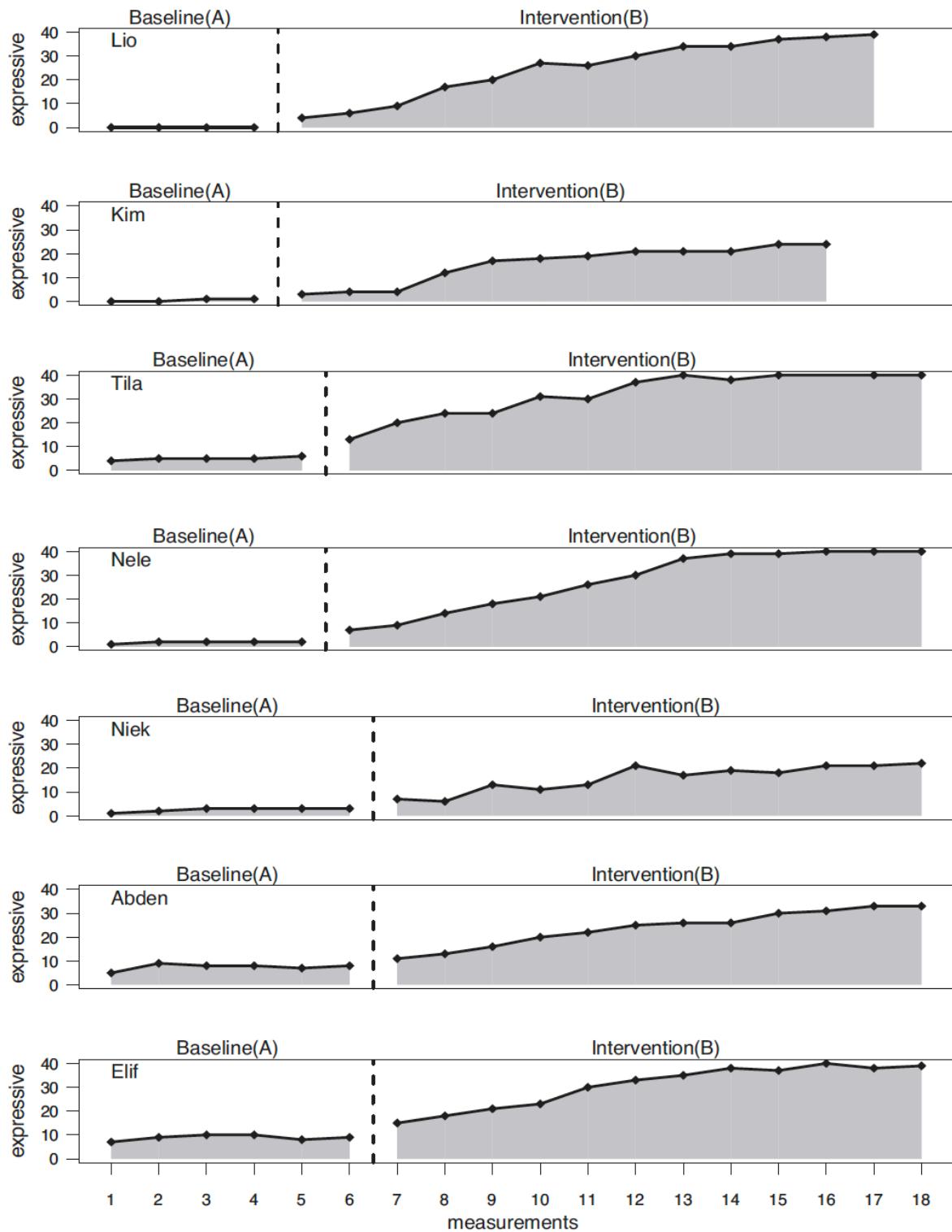
Figure 4

Letter sound fluency (LSF) in 1 min
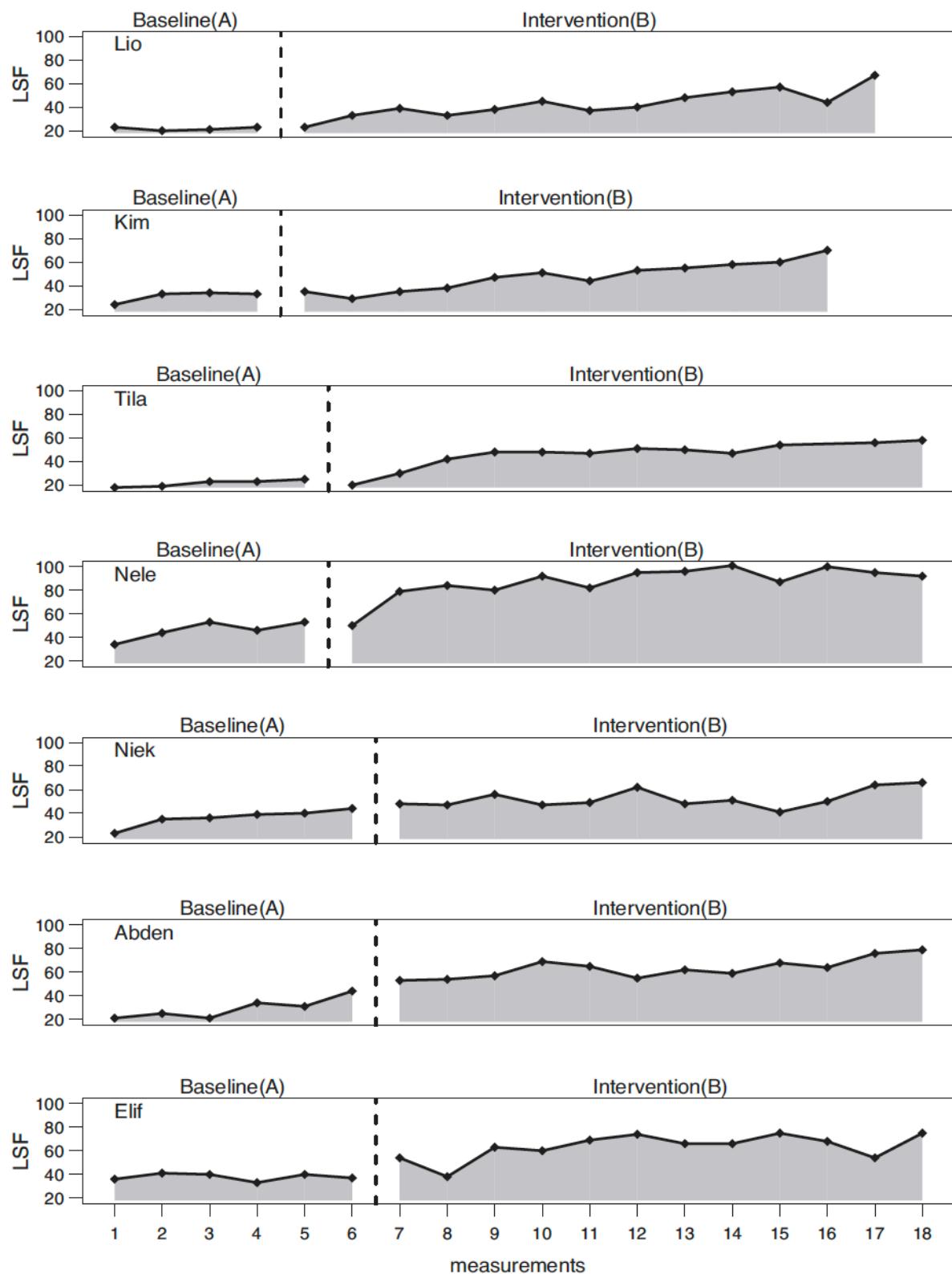
Figure 5

Amount of correctly read sight words
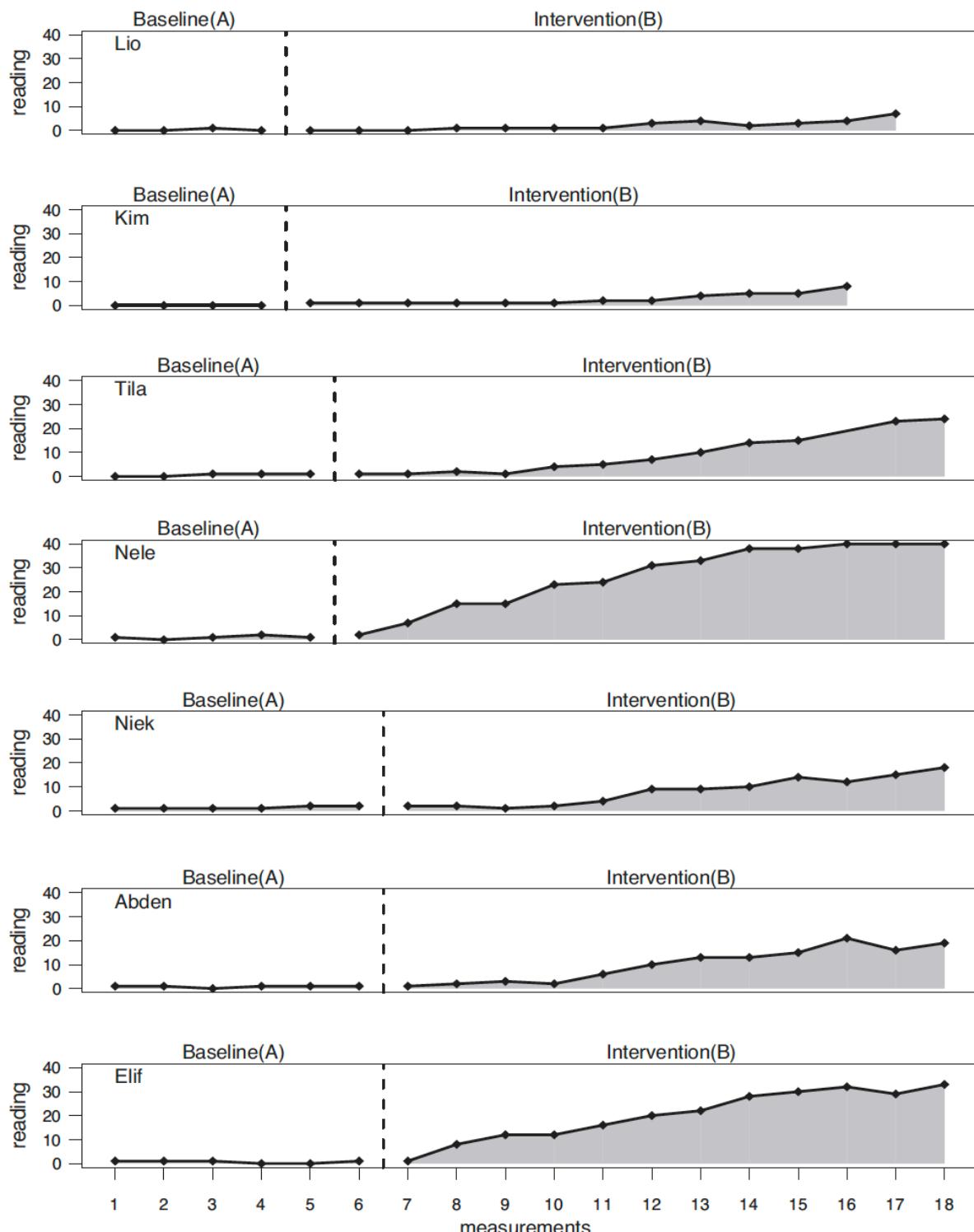
Results

Expressive Vocabulary

Overall, the visual baseline is very flat for all participants and there is a steady increase in the B phase. Tila ($M = 5.00$), Abden ($M = 7.50$), and Elif ($M = 8.83$) start with slightly higher values in the A phase while Lio ($M = 0.00$), Kim ($M = 0.50$), Nele ($M = 1.80$), and Niek ($M = 2.50$) start very low (see Table 2). The highest mean value in the B phase is shown by Tila ($M = 31.42$), and the lowest value is found in Niek ($M = 15.75$). The highest increase is shown by Kim (3,034%) and Lio (2,469%), and the lowest increase is shown by Abden (217.73%) and Elif (246.32%). Tila, Nele, and Elif reach the maximum possible score of 40.00 in the B phase (see Figure 3).

Table 2

Descriptive data for expressive vocabulary

Participants	N(A)	N(B)	M(A) SD	M(B) SD	MBDi	Md A	Md B	Max A	Max B
Lio	4	14	0.00 (0.00)	24.69 (12.41)	2,469%	0.00	27.00	0.00	39.00
Kim	4	14	0.50 (0.58)	15.67 (7.91)	3,034%	0.50	18.50	1.00	24.00
Tila	5	13	5.00 (0.71)	31.42 (9.30)	528,4%	5.00	34.00	6.00	40.00
Nele	5	13	1.80 (0.45)	27.69 (12.61)	1,438,4%	2.00	30.00	2.00	40.00
Niek	6	12	2.50 (0.84)	15.75 (5.63)	527,6%	3.00	17.50	3.00	22.00
Abden	6	12	7.50 (1.38)	23.83 (7.57)	217,73%	8.00	25.50	9.00	33.00
Elif	6	12	8.83 (1.17)	30.58 (8.98)	246,32%	9.00	34.00	10.00	40.00

N, measurements; M, mean; SD, standard deviation; MBDi, mean baseline difference; Md, median; Max, maximum; A, A phase; and B, B phase.

With regard to the overlap measures, the NAP shows the maximum value of 100.00 across all subjects ($p < 0.001 - p < 0.01$). The same picture can be seen for the PEM and the PAND. The Tau-U also shows statistically significant values ($p < 0.001$) which can be interpreted as a large change for Kim (0.69), Tila (0.70), and Niek (0.74) and as a very large change for Lio (0.83), Elif (0.84), Nele (0.88), and Abden (0.89; see Table 3).

Table 3

Overlap indices for expressive vocabulary

Participant	NAP	*p*	PEM	PAND	TauU	*p*
Lio	100.00	<0.01	100.00	100.00	0.83	<0.001
Kim	100.00	<0.01	100.00	100.00	0.69	<0.001
Tila	100.00	<0.001	100.00	100.00	0.70	<0.001
Nele	100.00	<0.001	100.00	100.00	0.88	<0.001
Niek	100.00	<0.001	100.00	100.00	0.74	<0.001
Abden	100.00	<0.001	100.00	100.00	0.89	<0.001
Elif	100.00	<0.001	100.00	100.00	0.84	<0.001

NAP, non-overlapping of all pairs; PEM, percentage of data points exceeding the median; and PAND, the percentage of all non-overlapping data.

The results of the regression analysis reveal for group 1 a statistically significant slope effect from A phase to B phase ($p < 0.05$) with a beta coefficient of 2.464 and thus, an improvement by this value per intervention session. Group 2 shows a statistically significant level effect ($p < 0.01$) as well as a slope effect ($p < 0.01$) with an improvement of 2.379 per session. For group 2, a significant level effect ($p < 0.05$) and slope effect ($p < 0.001$) can also be seen with a beta coefficient of 1.668. As expected, a statistically significant level effect ($p < 0.01$) from the A phase to the B phase and a significant slope effect ($p < 0.001$) from the A phase to the B phase can be seen. The subjects managed to improve by 2.259 more expressive correctly conscious words per intervention session (see Table 4).

Table 4

Regression model for expressive vocabulary across all participants (level 2-analysis)

	B	SE	*t*	*p*
Group 1				
Intercept	−0.250	3.805	−0.66	0.95
Trend	0.200	1.170	0.171	0.87
Level	1.697	2.693	0.630	0.53
Slope	2.464	1.188	2.075	<0.05
Group 2				
Intercept	2.500	3.005	0.832	0.41
Trend	0.300	0.791	0.379	0.71
Level	7.231	2.437	2.966	<0.01
Slope	2.379	0.814	2.924	<0.01
Group 3				
Intercept	5.311	3.326	1.597	0.12
Trend	0.276	0.405	0.681	0.50
Level	3.784	1.611	2.349	<0.05
Slope	1.668	0.429	3.883	<0.001
Overall				
Intercept	3.456	2.229	1.550	0.12
Trend	0.140	0.402	0.349	0.73
Level	4.086	1.369	2.985	<0.01
Slope	2.259	0.417	5.415	<0.001

Letter Sound Fluency

Visually, it can be said that the baselines here are not so flat compared to the expressive vocabulary and that positive trends can be partially assumed. Lio ($M = 21.75$), Kim ($M = 31.00$), and Tila ($M = 21.60$) start relatively low and also show no trend tendency in the A phase (see Table 5). Nele ($M = 46.00$), Niek ($M = 36.17$), Abden ($M = 29.33$), and Elif ($M = 37.83$) start with slightly higher values and show a positive trend tendency. Overall, however, there is also a clear increase for each test person in the B phase (see Figure 4). The overlap measures showed strong effects (94.00–100.00) for all children in the NAP, which were also statistically significant ($p < 0.01 – p < 0.001$). The PEM shows a maximum value of 100.00 for Lio, Nele, Niek, and Abden and a value of 91.67 for Kim, Tila, and Elif. The PAND also shows that the intervention was highly effective for all subjects (91.18–100.00). The Tau-U, taking into account a possible A phase trend, shows a moderate effect for Niek (0.52; $p < 0.01$), and a large change for the remaining children (0.62–0.69; $p < 0.001$; see Table 6). Regression analysis

showed neither a significant level effect ($p = 0.50$) nor slope effect ($p = 0.38$) for group 1. The same can be said for group 2. Group 3, on the other hand, shows a statistically significant level effect from the A to the B phase ($p < 0.05$), but also a trend in the A phase ($p < 0.05$). Overall, there is a significant level effect ($p < 0.05$) and an A phase trend ($p < 0.01$; see Table 7).

Table 5

Descriptive data for LSF

Partici-pants	N(A)	N(B)	M(A) SD	M(B) SD	MBDi	Md A	Md B	Max A	Max B
Lio	4	14	21.75 (1.50)	42.85 (11.50)	97,01%	22.00	40.00	23.00	67.00
Kim	4	14	31.00 (4.69)	47.92 (12.16)	54,58%	33.00	49.00	34.00	70.00
Tila	5	13	21.60 (2.97)	45.92 (10.90)	112,59%	23.00	48.00	25.00	58.00
Nele	5	13	46.00 (7.84)	87.50 (13.35)	90,22%	46.00	92.00	53.00	101.00
Niek	6	12	36.17 (7.19)	52.42 (7.81)	44,93%	37.50	49.50	44.00	66.00
Abden	6	12	29.33 (8.91)	63.41 (8.46)	116,20%	28.00	63.00	44.00	79.00
Elif	6	12	37.83 (3.06)	63.50 (10.79)	67,86%	38.50	66.00	41.00	75.00

N, measurements; M, mean; SD, standard deviation; MBDi, mean baseline difference (MBDi); Md, median; Max, maximum; A, A phase; and B, B phase.

Table 6

Overlap indices for LSF

Participant	NAP	p	PEM	PAND	TauU	p
Lio	98.00	<0.01	100.00	91.18	0.69	<0.001
Kim	94.00	<0.01	91.67	93.75	0.64	<0.001
Tila	95.00	<0.01	91.67	95.00	0.62	<0.001
Nele	97.00	<0.001	100.00	96.92	0.64	<0.001
Niek	99.00	<0.001	100.00	98.61	0.52	<0.01
Abden	100.00	<0.001	100.00	100.00	0.67	<0.001
Elif	96.00	<0.01	91.67	95.83	0.64	<0.001

NAP, non-overlapping of all pairs; PEM, percentage of data points exceeding the median; and PAND, the percentage of all non-overlapping data.

Table 7

Regression model for LSF across all participants (level 2-analysis)

	B	**SE**	**t**	**p**
Group 1				
Intercept	22.750	4.911	4.633	<0.001
Trend	1.450	1.529	0.949	0.35
Level	−2.410	3.520	−0.685	0.50
Slope	1.401	1.552	0.902	0.38
Group 2				
Intercept	25.100	14.713	1.706	0.10
Trend	2.900	1.838	1.578	0.13
Level	10.240	5.663	1.808	0.08
Slope	−0.442	1.890	−0.234	0.82
Group 3				
Intercept	25.664	4.480	5.724	<0.001
Trend	2.514	1.068	2.354	<0.05
Level	9.532	4.244	2.246	<0.05
Slope	−1.050	1.131	−0.928	0.36
Overall				
Intercept	23.614	5.182	4.557	<0.001
Trend	2.666	0.807	3.304	<0.01
Level	5.668	2.742	2.067	<0.05
Slope	−0.470	0.838	−0.561	0.58

Reading

Visual inspection shows enormously flat baselines with no positive trends. Significant increases in the B phases can only be found for five children. Lio and Kim initially reveal no improvement until the end, when there is a discrete increase. Kim ($M = 0.00$) and Lio ($M = 0.25$) start with the lowest values in the A phase and Niek ($M = 1.33$) and Abden ($M = 0.83$) with the highest values. The highest mean values in the B phase are shown by Nele ($M = 26.62$) and Elif ($M = 20.25$) and the lowest values by Lio ($M = 2.08$) and Kim ($M = 2.67$; see Table 8). The largest increase from A to B phase is observed in Nele (2562%) and Elif (2922%), and the least increase can be seen in Kim (267%) and Lio (732%). Only Nele reaches the maximum value of 40.00 in the B phase (see Figure 5). Lio and Kim show the lowest values with a maximum of 7.00–8.00. Further, the overlap measures for the NAP show a medium effect for Lio

(82.00; p < 0.05), Niek (90.00; p < 0.01), and Tila (92.00; p < 0.01) and a strong effect for Abden (97.00; p < 0.001), Elif (97.00; p < 0.001), Nele (99.00; p < 0.011), and Kim (100.00; p < 0.01). The PAND testifies medium effects for all except Nele and Kim, who show strong effects. A similar picture emerges for the PEM. The Tau-U displays a large change for Lio (0.61; p < 0.001), Kim (0.63; p < 0.001), Tila (0.69; p < 0.01), and Niek (0.69; p < 0.001). Abden (0.81; p < 0.001), Elif (0.87; p < 0.001), and Nele (0.88; p < 0.001) show a large to very large change (see Table 9). The results of the regression analysis at level 2 reveal no statistically significant level (p = 0.11) or slope effect (p = 0.18) for group 1. Group 2 shows a statistically significant slope from A to B phase (p < 0.05) with an increase of 2,503 correct words per intervention session. Group 3 indicates a very similar picture (slope; B = 2.502, p < 0.05). Overall, a significant slope effect can be observed with a beta coefficient of 1.224 (p < 0.05; see Table 10). Social Validity In terms of social validity, all participants have a very positive attitude towards the intervention overall (see Table 11). With regard to word reading, only Lio and Kim stated "partly agree." Overall, "completely agree" dominates on all items. The children found that the storytelling helped them, they understood the meaning of the promotion and would like to do more storytelling in school. The students also liked the self-graphing. Only Niek rated "partly agree."

Table 8

Descriptive data for words read correctly

Partici-pants	N(A)	N(B)	M(A) SD	M(B) SD	MBDi	Md A	Md B	Max A	Max B
Lio	4	14	0.25 (0.50)	2.08 (2.06)	732%	0.00	1.00	1.00	7.00
Kim	4	14	0.00 (0.00)	2.67 (2.31)	267%	0.00	1.50	0.00	8.00
Tila	5	13	0.60 (0.55)	8.92 (8.36)	1,386,67%	1.00	6.00	1.00	24.00
Nele	5	13	1.00 (0.71)	26.62 (13.35)	2,562%	1.00	9.00	2.00	40.00
Niek	6	12	1.33 (0.52)	8.17 (5.87)	514,29%	1.00	9.00	2.00	18.00
Abden	6	12	0.83 (0.41)	10.08 (7.10)	1,114,46%	1.00	11.50	1.00	21.00
Elif	6	12	0.67 (0.52)	20.25 (10.49)	2,922,39%	1.00	21.00	1.00	33.00

N, measurements; M, mean; SD, standard deviation; MBDi, mean baseline difference; Md, median; Max, maximum; A, A phase; and B, B phase.

Table 9

Overlap indices for correctly words read correctly

Participant	NAP	p	PEM	PAND	TauU	p
Lio	82.00	<0.05	76.92	70.59	0.61	<0.001
Kim	100.00	<0.01	100.00	100.00	0.63	<0.001
Tila	92.00	<0.01	75.00	82.35	0.69	<0.001
Nele	99.00	<0.001	100.00	94.44	0.88	<0.001
Niek	90.00	<0.01	91.67	75.00	0.69	<0.001
Abden	97.00	<0.001	91.67	83.33	0.81	<0.001
Elif	97.00	<0.001	91.67	86.11	0.87	<0.001

NAP, non-overlapping of all pairs; PEM, percentage of data points exceeding the median; and PAND, the percentage of all non-overlapping data.

Table 10

Regression model for words read correctly across all participants (level 2-analysis)

	B	SE	t	p
Group 1				
Intercept	0.000	0.892	0.000	1.00
Trend	0.050	0.323	0.155	0.88
Level	−1.226	0.743	−1.650	0.11
Slope	0.451	0.328	1.378	0.18
Group 2				
Intercept	0.050	5.937	0.008	0.99
Trend	0.250	1.183	0.211	0.83
Level	−1.226	0.743	−1.650	0.54
Slope	2.503	1.217	2.057	<0.05
Group 3				
Intercept	0.778	2.819	0.276	0.78
Trend	0.250	1.183	0.211	0.93
Level	−1.836	2.012	−0.913	0.54
Slope	2.503	1.217	2.057	<0.05
Overall				
Intercept	−1.603	2.911	−0.551	0.58
Trend	0.608	0.558	1.091	0.28
Level	−2.467	1.903	−1.296	0.20
Slope	1.224	0.579	2.114	<0.05

Table 11

Results of social validity questionnaire

Items	Lio	Kim	Tila	Nele	Niek	Abden	Elif
Storytelling helped me to be able to read words correctly	2	2	3	4	4	4	4
Storytelling helped me learn words and their meanings	3	4	3	4	4	3	4
Storytelling helped me to pro-nounce sounds correctly	4	4	4	3	4	4	4
I understood well the meaning of the promotion	4	4	4	4	4	3	3
I have learned a lot during story-telling	3	3	4	4	4	4	4
I gladly came to the intervention sessions	4	4	4	4	4	4	4
The self-graphing sheets were fun	3	3	4	4	2	4	4
The stories were great	4	4	3	4	4	4	3
I would like to do more with sto-ries in school	3	4	4	4	4	4	3

 0 = completely not agree; 1 = not agree; 2 = partly agree; 3 = agree; and 4 = completely agree.

Discussion

Main Findings

The study presented was designed to estimate the effects of a storytelling intervention on the variables: Vocabulary, LSF, and sight word reading in students with German as a second language with and without problem behavior. The background is the increasing number of students with GL2 and at the same time the increase of students with GL2 and weak school performance especially in the area of reading. L2 students are educationally disadvantaged due to their deficits in the language. It is of particular importance to teach these students the language adequately in a motivating way.

Overall, the results are consistent with findings that have looked at multi-component intervention (Foorman and Torgesen, 2001; Donegan and Wanzek, 2021) and the DCT (Paivio, 2008) which states that using verbal and non-verbal system of process information is highly effective in order to finally store information. Moreover, the findings are also consistent with the meta-analysis by Marulis and Neuman (2010) that conveying knowledge explicitly and implicitly in combination leads to the highest effects. Looking at the effectiveness on vocabulary acquisition, it can be seen that all subjects show an immense increase in the B phase, with all baselines being relatively flat. Niek, Kim, and Tila display the weakest effects, although

even these can be classified as large. Kim is by far the weakest in the vocabulary pretest with a PR of five. For her, this may be due to the fact that she has great problems building vocabulary overall. In contrast, Tila and Niek perform better in the vocabulary pretest, but unlike Kim, they have greater problems in LRB and the highest problem scores overall in the group. Particularly, problems in attention processing might be a reason here as describe in the literature (Peterson et al., 2013). Abden and Nele are among the strongest performers in terms of vocabulary, but both also show the best results in the vocabulary pretest. It might be easier for them to learn new words if their overall vocabulary is already larger. While Abden has problems with learning-related behavior, which does not seem to play a major role here, Nele shows no problems in this regard. The results of vocabulary acquisition are consistent with the findings of Barwasser et al. (2020, 2021), and Knaak et al. (2021).

Furthermore, for the second dependent variable LSF, the baseline results are higher, i.e., some children have already had experience with German letter sounds, while others reveal a flat baseline with lower values. Niek is the weakest and Lio and Abden the strongest. Niek shows by far the weakest results in the pretest in the area of sound categorization, which could be a reason for his problems in the area of LSF. Overall, Abden is also one of the weakest students in the phonological awareness pretest but sound categorization is his best sub-category with a PR of 28. Like Abden, Lio also has problems in learning behavior which also does not seem to play a major role. However, overall results indicate that the intervention does have a positive impact on LSF which is an important finding since Hulme et al. (2012) has shown that problems in LSF are related to later word-reading difficulties which are referred to almost the same age as the participants of the current study.

With regard to sight word reading, the overall performance is weaker, especially for Lio and Kim. Except for Nele and Elif, the others seem to take longer to automate the words. One explanation for this could be that less-proficient readers often take the non-lexical route because they have greater problems with the lexical route (De Jong et al., 2012). Thus, the children try to decode the words each time instead of storing them as a whole, for which the one second in the measurement is not sufficient. Thus, for these children it takes a longer time until they seem to change the route. Nele and Elif both have much higher scores in phonological awareness and also in pseudoword reading, which should make it easier for them to memorize the words as a whole more quickly, as they are better readers. In reading, they are among the strongest of the subjects in the pretests, along with Abden, who scores third best in overlap-indices. Elif, like Nele, shows no problems in learning-related behavior. Lio and Kim are among the weakest subjects overall in terms of reading and phonological awareness. Perhaps,

the Polish L1 also plays a role because L1 background can influence L2 word recognition (Wang and Koda, 2007). According to Catts (1993), phonological awareness is more closely related to word recognition than measures of vocabulary in young first grade children with phonological difficulties and Lio and Kim perform poorly in both areas. Another explanation could be that Lio and Kim might have problems in rapid automatized naming, which is important with regard to naming speed and the retrieval of sight words from the mental lexicon, especially in the German language (Landerl and Wimmer, 2008; Huschka et al., 2021). Nevertheless, Abden, Niek, and Tila also display severe problems in phonological awareness and need longer time to respond to the intervention in word recognition. Niek and Abden have better reading performance in the pretest while Tila performs similarly weak in the pretest as Lio and Kim. So, what could be the reason? In the case of Tila, it could actually be the learning-related behavior that causes problems, or frustration, while in the case of Niek and Abden, the behavioral problems do not seem to have such an impact. One reason could be the overall better reading performance of the two students, which counteracts the problem behavior.

Overall, the intervention seems to work really well for one variable and well for the other two. Storytelling seems to also have an effect on the reading of sight words and goes partly in line by meta-analytic finding by Roberts et al. (2020) who were focused on foundational reading instructions for students with problem behavior in grades K-12 ($g = 0.86$) as well as small group reading instruction for grade 1–4 (e.g., Scammacca et al., 2015). With regard to students who struggled with sight word reading, one can see that even with those displaying slow increase, the improvement seems to come after some time. Another assumption could be that the intervention should have been prolonged in order to achieve greater effects. Also, behavior might have played a role in some cases in combination with very low score in the pretesting. Reflecting on the importance of motivation, especially in language acquisition, self-graphing probably contributed in part to the effects, as studies have pointed to the effectiveness of self-graphing in intervention and especially in reading interventions (Stotz et al., 2008; Bowman-Perrott et al., 2013; Leko, 2016; Guzman et al., 2018; McKenna and Bettini, 2018). Especially regarding the social validity results where all children despite Niek, who seemed to be unsure, rated self-graphing as positive. Moreover, the results of the social validity questionnaire revealed that all participants rated the interventions as positive. With regard to reading, Lia and Kim gave worse scores than the others, but this is also understandable, since both could hardly benefit in sight word reading, also compared to the others.

It is also noticeable that the language background does not necessarily play a role. The Polish background is only noticeable when improving the visual vocabulary, but this does not

necessarily mean anything. The sample is too small to be able to make statements about this. Also, problem behavior did not seem to play a role across the board. This may be due to the fact that the children were taken out of the classroom and trained intensively in a small group. In general, small group interventions, especially with regard to reading, have been shown in a meta-analysis by Hall and Burns (2018) to achieve a large effect size for elementary students ($g = 0.64$; also see Nielsen and Friesen, 2012) which can be also referred to Roberts et al. (2020) who examined the effects for primary school students with behavioral problems in a meta-analysis.

Limitations

In addition to the promising results, there are some imitations: First, the intervention took place during the COVID-19 pandemic, where everyone in the school had to abide by specific rules and it was generally unruly in the school. Groups were therefore not allowed to be mixed from different classes. With regard to reading, it can be seen that those with very weak performance at the phonological level also have greater problems storing the words as sight words. Here, it would probably make sense to stay one level lower and train the LSF and other aspects of decoding more intensively. Furthermore, this is a multiple baseline study, which means that we focused on individual students, making it difficult to generalize the results. Nevertheless, the results give important indications with regard to the support of struggling students with GL2 with and without behavioral difficulties. The advantage of a multiple baseline study is that it allows us to see individual learning trajectories and to find out specifically how the intervention is received by different students.

Another limitation is that there is a certain probability that the children have also become better through the repeated measurements each time after the sessions. We have tried to counteract this by randomizing the order of the items in each test, but we cannot exclude it for sure. However, since there are no trends in the baselines where only testing was done, it could be argued that the influence of testing was not too great. A further minor limitation is the measurement time point of the first group in the baseline, since across the board at least five measurement time points are always recommended. in each phase. After Kratochwill et al. (2013), however, at least three measurement time points are also sufficient to be able to make a statement. Due to time constraints, it was not possible to extend the baseline. And, as with all multi-component interventions, of course, one does not know which component worked for which parts. At the current time, it is not possible to say exactly to which parts the various components (such as self-graphing and implicit vs. explicit teaching) have had on the dependent variables. Since this intervention seems to work in this package, it is basically not the intention to examine

the individual parts separately, as the package is very easy and straightforward to implement in the classroom.

Implications

A first goal would be to estimate the storytelling intervention on a larger sample and make generalized statements. Furthermore, the intervention would be compared to other interventions in order to see which support option seem to be most effective in the area of language acquisition. In the course of this, one could also look at whether the method also works with a whole class or if it is limited to small groups. In the context of digitalization and especially the current school closures worldwide due to the COVID-19 pandemic, which has once again shown how important digital learning is in schools, the storytelling intervention could be digitalized and made available *via* apps or web-based tools.

The intervention in its current form was rated very positively, which gives us an indication that despite the overall good effects on all three dependent variables, the intervention is accepted across all participants. The further implication of this is to continue to conduct the social validity survey in future research to gain more insight into the overall acceptability of the intervention, which according to Briesch et al. (2013) is a necessity in intervention research. Last implications are the different languages and behavioral problems. It would be interesting to see whether the effects differ between children from different language backgrounds (Wang and Koda, 2007). In addition, one could also record the abilities in the surveyed variables in the L1 in order to identify possible correlations here. Furthermore, the study looked at children with learning-related behavior problems. A continuation would be to see if the intervention would also help with students with disruptive behavior, which is a big challenge for teachers today (Rosenberg and Jackman, 2003). Also, measuring rapid automatized naming beforehand would be interesting since it is linked to rapid word retrieval and reading, particularly in the German language which is more transparent than, e.g., English (Landerl and Wimmer, 2008).

Conclusion

It is enormously important to support struggling language learners in all components of a language in order to provide equal chances with respect to school and later job possibilities, especially to actively address the results of the PISA survey (OECD, 2019). Also, Morgan et al. (2015) showed that first graders with reading problems are more likely to show off task-behavior and general problem behavior in grade 3. Also considering the meta-analysis by Chow and Wehby (2018) on the negative relationship of language problems and behavioral difficulties, it is imperative to counteract this, particularly when students already display some kind of problem behavior. Also, one should consider the Matthew effect that stronger readers become

stronger and weaker readers become weaker particularly in the first years of school because they start to dislike reading (Stanovich, 1986). Thus, early prevention in school failure is really important, specifically for students with GL2 and those with additional problem behavior who struggle with reading. This storytelling approach should give teachers, educators, and researchers an indication of how an intervention in this area could look like which can train different areas of language at the same time and matches the concept of inclusion by Booth and Ainscow (2011) to integrate students with different competencies and characteristics as well as from different backgrounds.

References

Aschenbrenner, K.-H., Bauer, W., Gehring, S., Hazzouri, N., Holdorf, K., Jeuk, S., et al. (2016). *Deutsch als Zweitsprache in der Grundschule. 2nd Edn.* Landesinstitut für Schulentwicklung (LS). Available at: https://www.schule-bw.de/themen-und-impulse/migration-integration-bildung/fakten_kontakte_angebote.html/praxis_wissen/handreichungen/deutsch_als_zweitsprache.pdf (Accessed March 28, 2021).

Baker, D. L., Santoro, L., Biancarosa, G., Baker, S. K., Fien, H., and Otterstedt, J. (2020). Effects of a read aloud intervention on first grade student vocabulary, listening comprehension, and language proficiency. *Read Writ.* 33, 2697–2724. doi: 10.1007/s11145-020- 10060-2

Baker, S. K., Santoro, L. E., Chard, D. J., Fien, H., Park, Y., and Otterstedt, J. (2013). An evaluation of an explicit read aloud intervention taught in whole-classroom formats in first grade. *Elem. Sch. J.* 113, 331–358.

Barwasser, A., Knaak, T., and Grünke, M. (2020). The effects of a multicomponent storytelling intervention on the vocabulary recognition of struggling English as a foreign language learners with learning disabilities. *Insight Learn. Disabil.* 17, 35–53.

Barwasser, A., Lenz, B., and Grünke, M. (2021). A multimodal storytelling intervention for improving the reading and vocabulary skills of struggling German as a second language adolescents with learning and behavioral problems. *Insight Learn. Disabil.* 18, 29–51.

Becker-Mrotzek, M., Hentschel, B., Hippmann, K., and Linnemann, M. (2012). *Sprachförderung in deutschen Schulen – die Sicht der Lehrerinnen und Lehrer Ergebnisse einer Umfrage unter Lehrerinnen und Lehrern.* Mercator-Instituts für Sprachförderung und Deutsch als Zweitsprache, Universität zu Köln. Available at: https://www.mercator-institut-sprachfoerderung.de/fileadmin/user_upload/Lehrerumfrage_Langfassung_final_30_05_03.pdf (Accessed March 20, 2021).

Booth, T., and Ainscow, M. (2011). *Index for Inclusion Developing Learning and Participation in Schools.* Bristol: Centre for Studies on Inclusive Education.

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. doi: 10.1080/02796015.2013.12087490

Brewster, J., Ellis, G., and Girard, D. (2002). *The Primary English Teacher's Guide*. New Edn. England: Pearson Education.

Briesch, A. M., Chafouleas, S. M., Neugebauer, S. R., and Riley-Tillman, T. C. (2013). Assessing influences on intervention implementation: revision of the usage rating profile-intervention. *J. Sch. Psychol.* 51, 81–96. doi: 10.1016/j. jsp.2012.08.006

Brysbaert, M., Mandera, P., and Keuleers, E. (2018). The word frequency effect in word processing: an updated review. *Curr. Dir. Psychol. Sci.* 27, 45–50. doi: 10.1177/0963721417727521

Catts, H. (1993). The relationship between speech-language impairments and reading disabilities. *J. Speech Hear. Res.* 36, 948–958.

Choo, L. B., Lin, D. T., and Pandian, A. (2012). Language learning approaches: a review of research on explicit and implicit learning in vocabulary acquisition. *Procedia Soc. Behav. Sci.* 55, 852–860. doi: 10.1016/j.sbspro.2012.09.572

Chow, J. C., and Wehby, J. H. (2018). Associations between language and problem behavior: a systematic review and correlational meta-analysis. *Educ. Psychol. Rev.* 30, 61–82. doi: 10.1007/s10648-016-9385-z

Clemens, N. H., Lai, M. H. C., Burke, M., and Wu, J. Y. (2017). Interrelations of growth in letter naming and sound fluency in kindergarten and implications for subsequent reading fluency. *Sch. Psychol. Rev.* 46, 272–287. doi: 10.17105/SPR-2017-0032.V46-3

Coltheart, M. (2005). "Modeling reading. The dual-route approach," in *The Science of Reading: A Handbook*. eds. M. J. Snowling and C. Hulme (New Jersey: Blackwell).

Curtis, P. R., Frey, J. R., Watson, C. D., Hampton, L. H., and Roberts, M. Y. (2018). Language disorders and problem behaviors: a meta-analysis. *Pediatrics* 142, e20173551. doi: 10.1542/peds.2017-3551

De Jong, C. G., Licht, R., Sergeant, J. A., and Oosterlaan, J. (2012). RD, ADHD, and their comorbidity from a dual route perspective. *Child Neuropsychol.* 18, 467–486. doi: 10.1080/09297049.2011.625354

DiGangi, S. A., Maag, J. W., and Rutherford, R. B. (1991). Self-graphing of on-task behavior: enhancing the reactive effects of self-monitoring on on-task behavior and academic

performance. *Learn. Disabil. Q.* 14, 221–230. doi: 10.2307/1510851

Donegan, E. E., and Wanzek, J. (2021). Effects of reading interventions implemented for upper elementary struggling readers. A look at recent research. *Reading Writing* 35, 62–71. doi: 10.1007/s11145-021-10123-y

Ehri, L. C. (2002). Phases of acquisition in learning to read words and implications for teaching. *Br. J. Educ. Psychol. Monogr. Ser.* 1, 7–28.

Ehri, L. C. (2005). "Development of sight word reading: phases and findings," in *Blackwell Handbooks of Developmental Psychology. The Science of Reading: A Handbook.* eds. M. J. Snowling and C. Hulme (New Jersey: Blackwell Publishing), 135–154.

Ender, A. (2016). And explicit cognitive processes in incidental vocabulary acquisition. *Appl. Linguis.* 37, 536–560. doi: 10.1093/applin/amu051

Federal Statistical Office (2020). *Ausländische Schüler/−innen nach Schularten.* Available at: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Schulen/Tabellen/allgemeinbildende-berufliche-auslaendische-schueler.html;jsessionid=2E832BA6153E455C10E6655AEF8A9F19.internet711?nn=209624 (Accessed October 20, 2020).

Fien, H., Santoro, L., Baker, S. K., Park, Y., Chard, D. J., Williams, S., et al. (2011). Enhancing teacher read alouds with small-group vocabulary instruction for students with low vocabulary in first-grade classrooms. *Sch. Psychol. Rev.* 40, 307–318. doi: 10.1080/02796015.2011.12087720

Foorman, B. R., and Torgesen, J. (2001). Critical elements of class-room and small-group instruction promote reading success in all children. *Learn. Disabil. Res. Pract.* 16, 203–212. doi: 10.1111/0938-8982.00020

Ghenghesh, P. (2010). The motivation of L2 learners: does it decrease with age? *Engl. Lang. Teach.* 3, 128–141.

Guzman, G., Goldberg, T. S., and Swanson, H. L. (2018). A meta-analysis of self-monitoring on reading performance of K-12 students. *Sch. Psychol. Q.* 33, 160–168. doi: 10.1037/spq0000199

Hall, M., and Burns, M. K. (2018). Meta-analysis of targeted small-group reading interventions. *J. Sch. Psychol.* 66, 54–66. doi: 10.1016/j.jsp.2017.11.002

Henrichs, J., Rescorla, L., Donkersloot, C., Schenk, J. J., Raat, H., Jaddoe, V. W., et al. (2013). Early vocabulary delay and behavioral/emotional problems in early childhood: the generation R study. *J. Speech Lang. Hear. Res.* 56, 553–566. doi: 10.1044/1092-4388(2012/11-0169)

Huang, X., Kim, N., and Christianson, K. (2019). Gesture and vocabulary learning in a second language. *Lang. Learn.* 69, 177–197. doi: 10.1111/lang.12326

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. H., and Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read. Combining intervention studies with mediation analyses. *Psychol. Sci.* 23, 572–577. doi: 10.1177/0956797611435921

Huschka, S. S., Georgiou, G. K., Brandenburg, J., Ehm, J. H., and Hasselhorn, M. (2021). Examining the contribution of RAN components to reading fluency, reading comprehension, and spelling in German. *Read. Writ.* 1–20. doi: 10.1007/s11145-021-10145-6

Hußmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., Lankes, E. M., et al. (2017). *IGLU 2016. Lesekompetenzen Von Grundschulkindern in Deutschland im Internationalen Vergleich*. Germany, New York, NY, USA: Münster.

Isbell, R., Sobol, J., Lindauer, L., and Lowrance, A. (2004). The effects of storytelling and story reading on the oral language complexity and story comprehension of young children. *Early Childhood Educ. J.* 32, 157–163. doi: 10.1023/B:ECEJ.0000048967.94189.a3

Jansen, R., Maljaars, J., Verhappen, A., Zink, I., Steyaert, J., and Noens, I. (2020). Problem behavior in young children referred with language difficulties: relations to language and intentional communication. *Autism Dev. Lang. Impairments* 5, 1–16. doi: 10.1177/2396941519900076

Jin, Z., and Webb, S. (2020). Incidental vocabulary learning through listening to teacher talk. *Mod. Lang. J.* 104, 550–566. doi: 10.1111/modl.12661

Kim, Y.-S. G., and Piper, B. (2019). Cross-language transfer of reading skills: an empirical investigation of bidirectionality and the influence of instructional environments. *Read. Writ.* 32, 839–871. doi: 10.1007/s11145-018-9889-7

Kitzinger, A. (2020). *METACOM. Symbolsystem zu Unterstützen Kommunikation*. Available at: https://www.metacom-symbole.de (Accessed February 05, 2021).

Knaak, T., Grünke, M., and Barwasser, A. (2021). Enhancing vocabulary recognition in English foreign learners with and without learning disability. The effects of a multi component storytelling intervention approach. *Learn. Disabilities: Contemp. J.* 19, 69–85.

Kramer, B., and McLean, S. (2019). L2 reading rate and word length. The necessity of character-based-measurement. *Read. Foreign Lang.* 31, 201–225.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., et al. (2013). Single-case intervention research design standards. *Remedial*

Spec. Educ. 34, 26–38. doi: 10.1177/0741932512452794

Landerl, K., and Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: an 8-year follow-up. *J. Educ. Psychol.* 100, 150–161. doi: 10.1037/0022-0663.100.1.150

Lane, J. D., Ledford, J. R., and Gast, D. L. (2017). Single-case experimental design: current standards and applications in occupational therapy. *Am. J. Occup. Ther.* 71, 7102300010. doi: 10.5014/ajot.2017.022210

Laurice, J. M., and Elisha, L. (2011). A review of the effects of self-monitoring on reading performance of students with disabilities. *J. Spec. Educ.* 45, 43–53. doi: 10.1177/0022466909349145

Leko, M. M. (2016). *Word Study in the Inclusive Secondary Classroom*. New York: Teacher College Press.

Lenhart, J., Lenhard, W., Vaahtoranta, E., and Suggate, S. (2018). Incidental vocabulary acquisition from listening to stories: a comparison between readaloud and free storytelling approaches. *Educ. Psychol.* 38, 596–616. doi: 10.1080/01443410.2017.1363377

Lervåg, A., and Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *J. Child Psychol. Psychiatry* 51, 612–620. doi: 10.1111/j.1469-7610.2009.02185.x

Ma, H. H. (2006). An alternative method for quantitative synthesis of single subject researches: percentage of data points exceeding the median. *Behav. Modif.* 30, 598–617. doi: 10.1177/0145445504272974

Manolov, R., Arnau, J., Solanas, A., and Bon, R. (2010). Regression-based techniques for statistical decision making in single-case designs. *Psicothema* 22, 1026–1032.

Marulis, L. M., and Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: a meta-analysis. *Rev. Educ. Res.* 80, 300–335. doi: 10.3102/0034654310377087

McClelland, M., Acock, A., and Morrison, F. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Child. Res. Q.* 21, 471–490. doi: 10.1016/j.ecresq.2006.09.003

McKenna, J. W., and Bettini, E. (2018). Improving reading fluency skills for secondary students with emotional and behavioral disorders. *Beyond Behavior* 27, 74–81. doi: 10.1177/1074295618779374

Mello, R. (2001). Building bridges: how storytelling influence teachers and student relationships. *J. High. Educ.* 17, 301–307.

Moll, K., and 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.

Morgan, P. L., Farkas, G., Tufis, P. A., and Sperling, R. A. (2015). Are reading and behavior problems risk factors for each other? *J. Learn. Disabil.* 41, 417–436. doi: 10.1177/0022219408321123

Nielsen, D. C., and Friesen, L. D. (2012). A study of the effectiveness of a small- group intervention on the vocabulary and narrative development of at-risk kindergarten children. *Read. Psychol.* 33, 269–299. doi: 10.1080/02702711.2010.508671

O'Brien, S., and Repp, A. C. (1990). Reinforcement-based reductive procedures: a review of 20 years of their use with persons with severe or profound retardation. *J. Assoc. Persons Sev. Handicap.* 15, 148–159.

Organisation for Economic Cooperation and Development (OECD) (2019). Programme for International Student Assessment (PISA) results from PISA 2018. Available at: https://www.oecd.org/pisa/publications/PISA2018_CN_DEU.pdf (Accessed March 03, 2021).

Paivio, A. (2007). *Mind and Its Evolution: A Dual Coding Theoretical Approach*. Mahwah, NJ: Erlbaum.

Paivio, A. (2008). "The dual coding theory," in *Educating the Other America*. ed. S. B. Neuman (Baltimore: Paul H. Brookes Publishing Company), 227–242.

Parker, R. I., Hagan-Burke, S., and Vannest, K. (2007). Percent of all nonoverlapping data (PAND): an alternative to PND. *J. Spec. Educ.* 40, 194–204. doi: 10.1177/00224669070400040101

Parker, R. I., Vannest, K. J., and Davis, J. L. (2011a). Effect size in single case research: a review of nine non-overlap techniques. *Behav. Modif.* 35, 302–322. doi: 10.1177/0145445511399147

Parker, R. I., Vannest, K. J., Davis, J. L., and Sauber, S. B. (2011b). Combining nonoverlap and trend for single-case research: tau-U. *Behav. Ther.* 42, 284–299. doi: 10.1016/j.beth.2010.08.006

Peterson, I. T., Bates, J. E., D'Onofrio, B. M., Coyne, C. A., Lansford, J. E., Dodge, K. A., et al. (2013). *J. Abnorm. Psychol.* 122, 542–557. doi: 10.1037/a0031963

Piasta, S. B., and Wagner, R. K. (2010). Developing early literacy skills: a meta-analysis of alphabet learning and instruction. *Read. Res. Q.* 45, 8–38. doi: 10.1598/RRQ.45.1.2

Raudenbush, S. W., and Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data*

Analysis Methods. Newbury Park, CA: Sage.

Reed, S. K. (2010). *Cognition: Theories and Applications. 8th Edn*. Bellmontt, CA, USA: Wadsworth, Cengage Learning.

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

Rosenberg, M. S., and Jackman, L. A. (2003). Development, implementation, and sustainability of comprehensive school-wide behavior management systems. *Interv. Sch. Clin.* 39, 10–21. doi: 10.1177/10534512030390010201

Scammacca, N. K., Roberts, G., Vaughn, S., and Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *J. Learn. Disabil.* 48, 369–390. doi: 10.1177/0022219413504995

Schmitt, N. (2014). Size and depth of vocabulary knowledge: what the research shows. *Lang. Learn.* 64, 913–951. doi: 10.1111/lang.12077

Schmitt, N., Jiang, X., and Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Mod. Lang. J.* 95, 26–43. doi: 10.1111/j.1540-4781.2011.01146.x

Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., and Kliegl, R. (2015). Child Lex–Eine lexikalische Datenbank zur Schriftsprache für kinder im Deutschen [childLex-A lexical database of written language for children in German]. *Psychol. Rundsch.* 66, 155–165. doi: 10.1026/0033-3042/a000275

Stanovich, K. (1986). Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Read. Res. Q.* 11, 360–407.

Stock, C., Marx, P., and Schneider, W. (2017). *Basiskompetenzen für Lese- Rechtschreibleistunge. Din Test zur Erfassung der Phonologischen Bewusstheit vom Ersten bis Vierten Grundschuljahr*. Göttingen: Hogrefe.

Stotz, K. E., Itoi, M., Konrad, M., and Alber-Morgan, S. R. (2008). Effects of self-graphing on written expression of fourth grade students with high-incidence disabilities. *J. Behav. Educ.* 17, 172–186. doi: 10.1007/s10864-007-9055-9

Suggate, S. P., Lenhard, W., Neudecker, E., and Schneider, W. (2013). Incidental vocabulary acquisition from stories: second and fourth graders learn more from listening than reading. *First Lang.* 33, 551–571. doi: 10.1177/0142723713503144

Swanson, E., Wanzek, J., Petscher, Y., Vaughn, S., Heckert, J., Cavanaugh, C., et al.

(2011). A synthesis of read-aloud interventions on early reading outcomes among preschool through third graders at risk for reading difficulties. *J. Learn. Disabil.* 44, 258–275. doi: 10.1177/0022219410 378444

Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., et al. (2016). The Single-Case Reporting guideline In BEhavioural interventions (SCRIBE) 2016 statement. *J. Sch. Psychol.* 56, 133–142. doi: 10.2522/ptj.2016.96.7.e1

Volpe, R. J., Casale, G., Mohiyeddini, C., Grosche, M., Hennemann, T., Briesch, A., et al. (2018). A universal behavioral screener linked to personalized classroom interventions: psychometric characteristics in a large sample of German school children. *J. Sch. Psychol.* 66, 25–40. doi: 10.1016/j.jsp.2017.11.003

Waddell, D. E., Nassar, S. L., and Gustafson, S. (2011). Single-case design in psychophysiological research: part II - statistical analytic approaches. *J. Neurother.* 15, 160–169. doi: 10.1080/10874208.2011.570693

Wajnryb, R. (2003). *Stories. Narrative Activities for the Language Classroom*. Cambridge: Cambridge University Press.

Wallace, C. (2014). *The Cambridge Guide to Teaching English to Speakers of Other Languages*. Cambridge, UK: Cambridge University Press.

Wang, M., and Koda, K. (2007). Commonalities and differences in word identification skills among learners of English as a second language. *Lang. Learn.* 55, 201–222. doi: 10.1111/j.0023-8333.2005.00290.x

Webb, S. A., and Chang, A. C.-S. (2012). Second language vocabulary growth. *RELC J.* 43, 113–126. doi: 10.1177/0033688212439367

Webb, S., and Nation, P. (2017). *How Vocabulary is Learned*. Oxford: Oxford University Press.

Weiß, R. H. (2007). *WS/ZF-R. Wortschatztest und Zahlenfolgentest – Revision. Ergänzungstest zum CFT 20-R. Manual*. Göttingen: Hogrefe.

Wilbert, J. (2021). Analyzing single-case data with R and SCAN. Available at: https://jazznbass.github.io/scan-Book/ (Accessed February 27, 2021).

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., and Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *J. Speech Lang. Hear. Res.* 50, 1093–1109. doi: 10.1044/1092-4388(2007/076)

Wright, A. (2013). *RBT: Creating Stories With Children*. Oxford: Oxford University Press.

Appendix E Article 5 *(peer-reviewed)*

Disclaimer for Dissertation Version:

This is the accepted manuscript of the article listed above. The final published version will be available from the publisher once released.

Abstract

Early language support is crucial for preventing educational inequality and promoting equal opportunities for all children. Phonological awareness and vocabulary are key predictors of both language acquisition and later reading and writing skills. In this single-case study, we investigated the effectiveness of early language support for preschoolers with German as a second language (L2), focusing on expressive vocabulary and grapheme-phoneme correspondence (GPC). Seven children (N = 7), aged five to six, with various first languages (L1), participated in a preschool in Germany. Two methods were compared: storytelling (incidental learning) and a combination of storytelling with flashcards (direct instruction (DI); intentional learning). Results showed minimal progress through storytelling alone, whereas the combined method led to significant improvements in vocabulary and GPC. Overlap metrics and visual analysis confirmed these findings. Social validity data reflected high acceptance among children and educators, despite some concerns about the intervention's scope. The study underscores the value of integrating implicit and explicit strategies in early language support, particularly for at-risk children, and advocates for structured language programs in preschool settings.

Keywords: Early Childhood Education, Storytelling, Direct Instruction, Grapheme-Phoneme-Correspondence, Vocabulary, L2 Acquisition, Learning Disabilities

Introduction

Preschool Language Development and Its Educational Relevance

Barriers to language acquisition often arise during early childhood. Without timely intervention, they can lead to significant long-term deficits in other areas of education. Many foundational language skills – particularly reading and writing – begin to develop early on, based on so-called precursor competencies (Ehri, 2005). These serve as the foundation for a comprehensive understanding of the language system. For this reason, early exposure to the target language is essential. Preschools provide an appropriate environment for this purpose.

Especially for families with limited educational resources or learning a second language (L2), early preschool attendance provides critical exposure to the target language and supports skill development. In the long term, this promotes educational success and improves the chances of achieving higher academic qualifications as well as better career prospects (Sachverständigenrat deutscher Stiftungen für Integration und Migration [SVR], 2013; Statistisches Bundesamt, 2024a).

However, in Germany, preschool attendance is not mandatory. While children with a migration background represent a growing proportion of the preschool population in Germany (43% of all under-sixes in 2023), only 76.8% of them attended preschool – compared to 99.3% of children without such a background (Statistisches Bundesamt, 2024a, 2024b, 2024c). Reasons for not utilizing preschool services include both normative beliefs and structural barriers – especially limited capacity (SVR, 2013). While participation in early childhood education does not guarantee positive outcomes, evidence suggests that high-quality early education settings can play a crucial role in supporting children's language development and social inclusion, particularly for children from multilingual or socioeconomically disadvantaged backgrounds (Becker & Lauterbach, 2016; Organisation for Economic Co-operation and Development [OECD], 2020).

Although a migrant background does not necessarily imply that a child acquires German as L2, statistics from 2022 indicate that only 51% of households with a migration background use German at home, and these households also tend to have fewer books and engage less frequently in shared reading activities (Anger et al., 2024; Lewalter et al., 2023).

These early language gaps are difficult to close later on, as highlighted by recent the Programme for International Student Assessment [PISA] and other studies' results in reading literacy (Institute for Quality Development in Education [IQB], 2022; OECD, 2023). Children with a migration background, in particular, tend to perform worse than their peers without a migration background (e.g., Valcárcel Jiménez et al., 2024).

Statistical surveys from the German capital Berlin from 2017 showed that among children with a migration background who either do not attend preschool or attend for less than six months, about 86% have poorer German skills at school entry compared to their native-speaking peers (Senatsverwaltung für Gesundheit, Pflege und Gleichstellung Berlin, 2018).

Currently, there are no binding regulations in Germany for early language support in preschools. While general educational guidelines exist, there is a lack of concrete requirements – similar to core curricula in schools – regarding specific language competencies. This absence leads to significant disparities in language development and, consequently, unequal starting conditions when children enter school.

Introducing clearly defined guidelines for implementing language support in early childhood could help reduce language barriers not only for children learning German as L2, but also for those at increased risk of learning disorders (Barwasser et al., 2025; Valcárcel Jiménez et al., 2024). Early intervention enables children with learning difficulties to catch up developmentally with their peers and to acquire long-term strategies that facilitate further learning (Barwasser et al., 2025). Without such support, these children may experience continued academic failure, resulting in educational disadvantages (Anger et al., 2024; Grünke & Bracht, 2025). Early intervention can play a crucial role in mitigating these disadvantages.

Children who struggle with phonological awareness or vocabulary development during preschool are at a significantly increased risk of later difficulties in fundamental academic skills like reading and writing (Karipidis et al., 2018). In Germany, migration background has been linked to lower phonological awareness and vocabulary levels in preschool children (Valcárcel Jiménez et al., 2024). Among phonological awareness components, phonemic awareness—the understanding that words consist of distinct sounds – and letter-sound knowledge (phonics) are the strongest predictors of literacy success in the first two school years (Carson et al., 2018). Carson et al. (2018) highlight the need to nurture both aspects early on and stress that research often overlooks narrowly defined phonological awareness, such as sound-level processing, which should receive greater focus in preschool settings. Further, early vocabulary knowledge serves as a foundational skill that directly impacts their ability to access and benefit from classroom instruction for L2 learners (Lesaux et al., 2007; Valcárcel Jiménez et al., 2024).Importantly, targeted early interventions can substantially reduce the risk of diagnosed learning disorders or severe learning delays (Barwasser et al., 2025; Grünke & Bracht, 2025).

Precursor Competencies in Language Development

Early language intervention should therefore focus on foundational skills – commonly referred to as precursor competencies. These include, in particular, phonological awareness,

which helps children recognize and process linguistic structures such as sounds and letters within a language's sound system, as well as the development of a broad vocabulary that supports both receptive and expressive language abilities (Barwasser et al., 2025; Grigorakis et al., 2022).

Grapheme-Phoneme Correspondence (GPC)

Grapheme-phoneme correspondence, a specific aspect of phonological awareness, refers to the accurate matching of letters to their respective sounds. This correspondence is language-dependent, as each language has its own phonological system, meaning that letters may be pronounced differently depending on the language (Bassetti at al., 2022; Sammour-Shehadeh et al., 2025). This poses a significant challenge especially for L2 learners, who often must adapt to an entirely new phonological system that may differ greatly from that of their L1 (Bassetti et al., 2022).

Vocabulary

Vocabulary is essential for both understanding and using language, as linguistic information is composed of a vast array of words. Receptive vocabulary is key to understanding language, while expressive vocabulary is crucial for the independent use of language. The larger a person's receptive and expressive vocabulary, the more extensive and varied their ability to express themselves linguistically (Dixon et al., 2022; Miralpeix & Muñoz, 2018; Yang et al., 2021). Previous studies have shown that receptive vocabulary typically develops before expressive vocabulary (Dixon et al., 2022). This means that using words independently is often more challenging for language learners than simply understanding them.

However, expressive vocabulary is vital for both everyday and academic language use, as it enables clear and nuanced verbal expression. For L2 learners, building expressive vocabulary is particularly difficult because it is very limited at the beginning of language acquisition and expands only gradually (Dixon et al., 2022; Nation, 2013). An insufficiently developed expressive vocabulary often leads to difficulties in other areas of linguistic competence, such as spoken and written expression as well as subject-specific language requirements (Dixon et al., 2022).

Approaches to Early Language Instruction

Incidental vs Intentional Learning

Research shows that children, in particular, acquire many aspects of phonological awareness and vocabulary through their environment and direct linguistic encounters (Nation, 2013). The learning and adaptation of language structures often occur incidentally during context-bound interactions with language in everyday life (Nation, 2013; Neumann et al., 2022;

Uchihara et al., 2019). This type of learning is referred to as incidental learning. In this process, knowledge is not acquired through explicit instruction, but rather through unintentional learning effects that arise from everyday language experiences.

Even without formal guidance, children are capable of developing essential language skills simply by listening and actively engaging in spoken interactions (Nation, 2013). A study by Neuman et al. (2022), for example, demonstrates that children are able to pick up new words from a variety of daily situations and integrate them into their vocabulary. This process of incidental learning is often described as fast mapping, where children connect new words and linguistic structures to what they already know, thereby continuously expanding their language abilities (Vlach & Sandhofer, 2012).

Children do not acquire language skills solely through incidental learning, but also through its counterpart – intentional learning (Baron & Arbel, 2022; Gallagher et al., 2019; Junttila & Ylinen, 2020). In intentional learning, attention is deliberately directed toward specific content with the conscious goal of acquiring knowledge. This type of learning is purposeful and can be supported through targeted instructional strategies.

GPC, in particular, typically requires explicit training, as it is often not adequately addressed in everyday language interactions (Baron & Arbel, 2022; Webber et al., 2024). Similarly, an expanded vocabulary that goes beyond common, everyday contexts can be systematically developed through intentional learning (Dixon et al., 2022; Junttila & Ylinen, 2020).

A study by Junttila and Ylinen (2020) examined the effectiveness of various learning approaches and found that while children were able to acquire new words through all methods, explicit instruction and focused attention on the learning content resulted in significantly better learning outcomes. Ultimately, both incidental and intentional learning have proven to be effective methods of language acquisition (de Vos et al., 2018). However, their effectiveness depends greatly on the individual linguistic background and developmental stage of the learner. For children learning German as a L2, both approaches are essential (de Vos et al., 2018).

On one hand, they benefit from language-rich environments – such as interactions with German-speaking peers or everyday contexts in preschool settings. On the other hand, studies like the ones conducted by Barwasser et al. (2021, 2022) demonstrate that targeted language support interventions in primary school can significantly reduce or even close existing gaps in L2 acquisition. In the studies, a multicomponent intervention – consisting of storytelling, explicit practice, and motivational strategies – led to significant improvements in both the receptive and expressive vocabulary and letter sound fluency (LSF) of students in grades three

through five. This finding highlights the importance of combining incidental and intentional learning, particularly for children acquiring German as L2 (Gallagher et al., 2019).

How to realize incidental and intentional learning?

Storytelling. Storytelling is a widely used didactic approach to support vocabulary development (Barwasser et al., 2021, 2022; Isbell et al., 2004). It involves expressive and emotionally engaging reading, often accompanied by gestures and facial expressions, to present new words in meaningful contexts. While new vocabulary is deliberately embedded in the narrative, children are not explicitly instructed to memorize it – making storytelling fundamentally a form of incidental learning. Through exposure to rich language in engaging stories, children intuitively acquire word meanings as part of the overall listening experience (Isbell et al., 2004). The effectiveness of storytelling has been demonstrated across various age groups and learner populations. Beyond vocabulary acquisition, it has also shown benefits in areas like cultural understanding and reading fluency (Barwasser et al., 2021, 2022; Isbell et al., 2004). A meta-analysis by Hostetter (2011) highlights that gestures and facial expressions significantly enhance language learning outcomes. Moreover, the motivating nature of stories - through personal relevance, appealing themes, or relatable characters—further supports their educational value (Flynn, 2004).

Flashcards. Studies by Barwasser et al. (2021, 2022) further emphasize the effectiveness of a combined instructional approach (intentional learning) that integrates storytelling with direct instruction (DI) – such as the use of flashcards. This combination has been shown to be particularly beneficial for reading development and vocabulary acquisition.

This approach is supported by the Dual Coding Theory (Paivio, 1991), which posits that learning is facilitated when multiple sensory channels – such as auditory and visual stimuli – are activated simultaneously. Using flashcards alongside storytelling enhances the clarity and repetition of learning content through visual reinforcement, making it more accessible for children. This multimodal approach can be especially effective in reinforcing GPC, as it links auditory cues (spoken words and sounds) with visual representations (letters and images), thereby strengthening foundational reading skills (Feng & Webb, 2020; Karipidis et al., 2018).

Purpose of this Study

While storytelling without DI aligns with the concept of incidental learning – as learning occurs through listening and contextual understanding – DI corresponds to intentional learning, where content is explicitly taught and practiced. This leads to a central question: to what extent is incidental learning already effective for children in the early stages of L2 acquisition, or does it need to be complemented by intentional learning to ensure effective outcomes?

Neumann et al. (2022) suggest that children learn extensively through contextual encounters with language. In contrast, another study by Law et al. (2019) indicates that direct linguistic instruction may be necessary, especially in the early phases of language acquisition. Against this backdrop, the following research questions arise for the present study:

1) To what extent does the use of storytelling with and without direct instruction (DI) influence the expressive vocabulary of preschool children learning German as a second language (L2)?

2) To what degree does storytelling with and without direct instruction affect grapheme-phoneme correspondence (GPC) in this target group?

3) How are the different instructional approaches (storytelling with and without direct instruction) perceived by the children themselves and by the educators?

Method

Participants and Setting

The study was conducted at a preschool in North Rhine-Westphalia, Germany, and involved three care groups. The preschool was selected after several institutions were contacted about participating in an early language support initiative. This preschool expressed interest and met key criteria: appropriate age group, sufficient group sizes, willingness to cooperate, and geographic proximity to the university. A total of twelve children, aged five to six, participated in the intervention program. All children were attending their last year of preschool, which in the German educational system precedes the start of primary school (Grade 1). One child, who took part in the intervention but was a native German speaker, was excluded from data collection due to the specific focus of the study. However, at the express request of the supervising educator and to ensure balanced group sizes, the child was included in the sessions. Inclusion criteria for the analysis required at least three baseline measurement points and participation in approximately three-quarters of the intervention sessions (a maximum of four missed sessions). Four children did not meet these criteria and were therefore excluded from the analysis.

The final sample consisted of seven children (four girls and three boys), all with a migration background and collectively representing four different L1: Turkish (four children), Russian, Arabic, and Kurdish (one child each) (Table 1). To protect their privacy, all participating children were assigned pseudonyms known only to the project team. The children's legal guardians were fully informed about the project and provided written consent for voluntary participation. All sessions took place during regular care hours in the familiar environment of the preschool facility.

Table 1

Demographic Characteristics of the Participants

		Aylin	Ivan	Defne	Alara	Tugce	Leith	Aras
Gender		female	male	female	female	female	male	male
Age		6	5	6	6	5	5	5
Migration		Yes	Yes	Yes	Yes	Yes	Yes	Yes
LFG		Yes	Yes	Yes	Yes	Yes	Yes	Yes
L1		Turkish	Russian	Turkish	Turkish	Turkish	Arabic	Kurdish
GPC		17.0	4.0	1.0	9.0	1.0	6.0	0.0
LSF		22.0	3.0	1.0	8.0	0.0	5.0	0.0
	SOR	87.9	16.4	35.8	87.9	35.8	1.8	6.7
	PS	87.5	27.1	16.8	16.4	27.1	2.8	8.1
	RR	46.2	62.5	8.9	46.5	19.1	0.8	0.8
TEPHOBE	ISC	87.2	23.2	11.4	87.2	11.4	3.7	3.7
	O	89.8	29.1	6.9	69.9	12.2	0.4	0.4
	RAN C	61.0	N.A.	49.6	2.6	10.3	0.9	1.3
	RAN O	68.2	12.9	29.0	5.6	7.1	0.4	7.3
ITRF-G		3.0	6.0	2.0	0.0	1.0	7.0	10.0

Note. Lingua Franca German (LFG); First Language (L1); Grapheme-Phoneme Correspondence (GPC); Letter Sound Fluency (LSF); TEPHOBE test, percentile ranks at Synthesis of Onset and Rime (SOR), Phoneme Synthesis (PS), Rhyme Recognition (RR), Initial Sound Categorization (ISC), Overall (O), Rapid Automatized Naming Colors (RAN C), Rapid Automatized Naming Objects (RAN O); Integrated Teacher Report Form German (ITRF-G), (cut-off ≥13) conspicuous in behaviour; Not Applicable (N.A.)

Initial testing revealed significant performance differences among the participating children (Table 1). Defne, Tugce, and Aras showed with very low scores in both letter-sound-fluency (LSF) and GPC, ranging from 0 to 1. In contrast, Aylin demonstrated relatively high initial proficiency, correctly identifying 17 sounds and producing 22 correct sounds within 60 seconds in the LSF assessment.

In the area of phonological awareness, as measured by the TEPHOBE, four children demonstrated clearly below-average performance, with percentile ranks ranging from 0.4 to 12.2. Two children performed within the average range, with percentile ranks of 29.1 and 69.9, respectively. Once again, Aylin achieved the highest results, scoring a percentile rank of 89.8, placing her well above average.

Substantial heterogeneity was also observed in the rapid naming subtests. Only Aylin and Defne achieved average performance, while all other children scored below the 15th percentile, indicating clearly below-average abilities. The color-naming subtest for rapid naming could not be administered to Ivan, as it became apparent during the practice items that he was

unable to name one of the tested colors, thus not meeting the prerequisites for testing (Mayer, 2020).

The testing and intervention sessions were carried out by three master's students enrolled in a teacher training program for special education. Prior to implementation, they received comprehensive training from the project leadership team on the diagnostic procedures and intervention protocols. Additionally, they were provided with a detailed guide outlining each step of the intervention process.

Measurements and Dependent Variables

Before the baseline phase began, initial assessments were conducted with the children to evaluate their current levels of vocabulary and phonological awareness. This preliminary assessment included both standardized and non-standardized pretests, as well as questionnaires completed by the educators.

Screenings

Integrated Teacher Report Form – German Language Version (ITRF-G). To assess learning-related and oppositional behavior in children, the ITRF questionnaire was filled out by the educational staff. For this study, a short German-language version of the original instrument (Volpe et al., 2018) was applied. Educators rated 16 statements per child on a four-point Likert scale (0 = no noticeable behavior to 3 = highly noticeable behavior). Eight items assessed learning-related behavior, while the other eight focused on oppositional behavior. A total score of 13 or higher is considered an indicator of problematic behavior (Volpe et al., 2018). In this sample, the total and subscale scores for all children were below the cut-off value, indicating that no child exhibited behavior considered problematic.

Test to Record Phonological Awareness and Rapid Automatized Naming (TE-PHOBE). The TEPHOBE is a standardized diagnostic tool designed to assess phonological awareness and rapid automatized naming ability in preschool children up to the second grade (Mayer, 2020). The subtests for phonological awareness in preschool include the following task areas: onset-rhyme synthesis, phoneme synthesis, rhyme recognition, and initial sound categorization. The subtests for rapid naming at the preschool level involve quick naming of objects and colors. The full test takes approximately one hour to administer and has been normed on both monolingual German-speaking and bilingual children (Mayer, 2020).

Letter Sound Fluency. To assess LSF, children were presented with a chart containing letters of the alphabet in print, displayed multiple times in randomized order. The letters <q>, <x>, <y> and <z> were excluded, as they cannot be clearly assigned to specific phonemes and have a low initial sound frequency in the German vocabulary. Children were asked to name the

letters solely by their corresponding sounds, not by their letter names, within a 60-second time limit. The number of correctly articulated sounds within this time frame was recorded as the raw score for this assessment.

Expressive Vocabulary Word List. To assess expressive vocabulary, picture cards using Metacom symbols (Kitzinger, 2023) were presented to the children. Based on the results, 30 words that none of the participating children were able to name correctly were identified. These words formed the basis of the target vocabulary for both the intervention and ongoing measurement throughout the study.

Dependent Variables

Grapheme-Phoneme Correspondence. GPC was measured both in the pretest and as a dependent variable using a PowerPoint presentation (PPT). This presentation displayed each letter of the alphabet once, in randomized order—excluding the letters <q>, <x>, <y> and <z>. Children were asked to produce the corresponding sound for each letter. The raw score was calculated as the number of correctly named sounds, with a maximum possible score of 22. A new randomization of letter order was used for each testing session.

Expressive Vocabulary. To measure expressive vocabulary (as a second dependent variable) again a PPT was used and the 30 target pictures were shown to the children using Metacom symbols at the end of each intervention session. On each PPT slide there was one picture. The order of the pictures varied at each measurement point. The raw score was based on the number of correctly named pictures, with a maximum of 30 points.

Design

This study was conducted within a single-case research framework, using a multiple baseline across participants design in an A-B-BC format (Ledford & Gast, 2024). The design included three consecutive phases: a baseline phase (A), a first intervention phase with storytelling (B), and a second intervention phase combining storytelling with DI (BC).

Within the multiple-baseline design, the twelve participating children were divided into three groups of four children each. Each group underwent a baseline phase of varying length (Kratochwill et al., 2013). As a result, Group 1 completed a baseline of five sessions, Group 2 of six sessions, and Group 3 of seven sessions. Accordingly, the duration of the two intervention phases also varied between the groups. This staggered approach increases internal validity, as it allows observed effects to be more clearly attributed to the specific intervention phases while better controlling for potential confounding variables, such as general developmental progress (Ledford & Gast, 2024). A stable baseline also helps minimize the Hawthorne effect,

which could otherwise arise merely from increased interaction with interventionists (Chiesa & Hobbs, 2008; Ledford & Gast, 2024).

The intervention was delivered three times per week, with a maximum of 15 sessions implemented for Group 1. The entire project spanned approximately eight weeks. Data were collected at a total of 21 measurement points, following each baseline and intervention session. All assessments were conducted in individual testing sessions.

Procedures

Phase A – Baseline

The baseline phase served to determine each child's individual starting level and to ensure internal validity (Ledford & Gast, 2024). During this phase, children were already working within their assigned small groups, but no instructional content was delivered. The duration of the baseline sessions matched that of the later intervention sessions in order to maintain comparable conditions across phases. Activities during this phase were unrelated to the target learning content and included tasks such as creative exercises and group games. At the end of each session, a measurement of the dependent variables was conducted.

Intervention Material (B and BC phase)

The intervention materials included content for both the storytelling-only phase (B) and the combined storytelling and DI phase (BC). A unique story was created for each session in both the B and BC phases, averaging around 600 words per story. The stories were printed in letter-sized format, using a legible font size, and inserted into a tabletop flip chart, allowing the children to visually follow along as the story was read.

Each story contained ten randomly selected vocabulary words from the total set of 30 target words. Across all stories, each of the 30 words appeared with equal frequency. Within a single story, each of the ten selected words occurred twice and was highlighted in color within the text (Figure 1). A total of 16 independent stories were created, each centered around an adventure involving the two main characters – a fox and an okapi. These two animals were represented by hand puppets during the sessions and could be incorporated into the play by the children as a motivational element (Flynn, 2004). In Phase BC, flashcards were additionally used. Each flashcard represented one of the 30 vocabulary words and was designed in letter-sized format, displaying both the written word and a colorful illustration using Metacom symbols.

Phase B – Intervention with Storytelling

The first intervention phase focused on the storytelling method. In each session, a story containing the predefined target vocabulary words was read aloud to the children. The children

sat in a semicircle in front of two interventionists, who presented the story using a presentation binder, incorporating expressive gestures, facial expressions, and a puppet representing an animal character from the story. The vocabulary words were emphasized through prosody but were not explicitly discussed or practiced with the children.

Phase BC – Intervention with Storytelling and Direct Instruction

In the second intervention phase, the storytelling method was enhanced with elements of DI. As in Phase B, the story was read aloud, but this time supplemented by the targeted use of flashcards. Each time a vocabulary word appeared in the story, the narration was briefly paused, and the corresponding word was visually presented using a flashcard containing both the written word and a related image (Figure 1).

Figure 1

Example of a Page within a Story and a Flashcard (Translated)

Pepsi says:

"I need my equipment. That's why I packed a magnifying glass, an **envelope**, a flashlight, gloves, sunglasses, hiking boots, and a clothes hanger.

With these things, I can solve all the puzzles we'll find on our way.

Here, hold the **flashlight** so we can see where we are walking and don't trip over a **branch** on the ground."

envelope
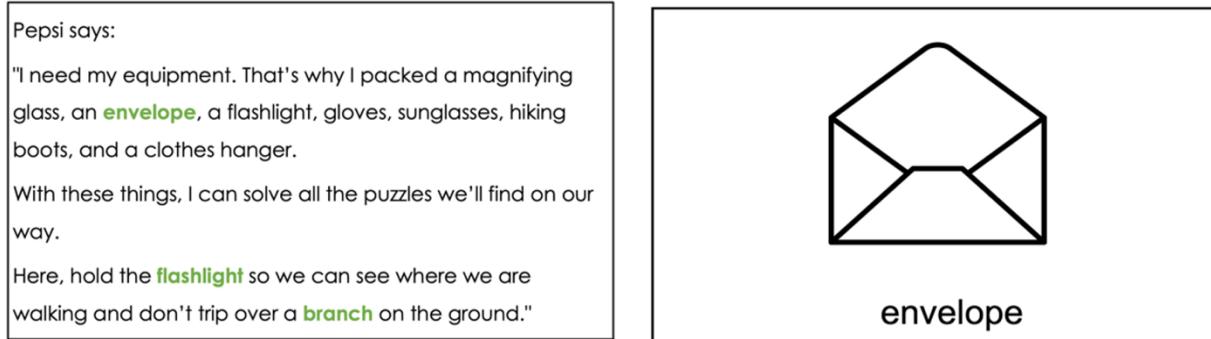
The interventionists then provided a brief explanation of the word's meaning and explicitly addressed its phonological structure. The individual sounds of the word were articulated separately and the children were prompted to repeat them. After this short instructional segment, the storytelling continued until the next vocabulary word appeared. This structured integration of auditory and visual learning pathways was designed to specifically support both vocabulary acquisition and GPC.

Social Validity

To assess the project's acceptance and emotional impact, social validity was evaluated from the perspectives of both the participating children and the supervising educators. The assessment of social validity is a key component in the evaluation of practice-oriented interventions, as it provides valuable insights into feasibility, subjective perception, and potential implications for educational practice (Kazdin, 1977).

At the end of the project, the participating children received a questionnaire containing a total of 13 items, thematically assigned to the categories of emotional response and motivation (e.g., "I enjoyed the sessions"), perceived learning (e.g., "I learned new words through the sessions"), and social aspects (e.g., "I liked working in a group"). For each statement, children were able to give a rating using visually supported response options: a green thumbs-up (2 points), a yellow sideways thumb (1 point), and a red thumbs-down (0 points). The educators assisted the children in understanding the questions but refrained from influencing their responses in any way.

The educational staff received a separate questionnaire consisting of six items, which were rated on a five-point Likert scale (0 = "strongly disagree" to 4 = "strongly agree"). The items addressed the perceived effectiveness of the method (e.g., "I feel that the children benefited from the intervention"), the children's enjoyment of participation (e.g., "I feel the children enjoyed the sessions"), and the method's applicability in everyday educational practice (e.g., "I would use the storytelling program in daily routines").

Treatment Fidelity

In addition, treatment fidelity was systematically assessed to ensure that the intervention was implemented as planned (Nelson et al., 2012; Sanetti et al., 2021). The items of the questionnaire assessed whether the intervention followed the planned timeline, whether there were any external disruptions or other notable incidents, and whether the procedure was carried out as usual. Approximately one-third of the intervention sessions were attended by external observers who were informed about the intervention procedures but were not involved in the diagnostic process or the actual implementation. Their role was to objectively monitor adherence to the intervention protocol. Furthermore, the interventionists completed a session log after each unit to document the implementation and note any particularities or deviations. The interrater reliability for treatment fidelity was 100%, indicating complete agreement among all observers regarding the correct implementation of the intervention (Wolery et al., 2011).

Data Analysis

Data analysis followed established quantitative and visual single-case analysis methods as outlined by Kratochwill et al. (2013). Statistical calculations were carried out using the Scan Package by Wilbert and Lüke (2025) in RStudio. The goal of the analysis was to clearly identify both the individual developmental trajectories of the participants and the phase-specific effects of the intervention. As part of the descriptive analysis, means (M), standard deviations (SD), and maximum values (Max) were calculated for each child and each phase. These metrics allow for an assessment of performance variability within and between the individual phases

(Ledford & Gast, 2024). In addition, statistical values for overlap measures were calculated. For all overlap indices, the significance level was determined and indicated using *p*-values. The values were calculated and analyzed for the transitions from Phase A to B as well as from B to BC.

The Tau-U accounts for both level changes and existing trends within phases and represents a robust measure, particularly in studies with small sample sizes (Parker et al., 2011). Values between 0.21 and 0.60 indicate a medium effect, between 0.61 and 0.80 a moderate effect, and from 0.81 onward a strong effect. Calculation was based on the formula: A vs. B + TrendB – TrendA. The Percentage of Data Exceeding the Median Trend (PEM-T*)* indicates how many data points within a phase exceed the median of the preceding phase, while also taking trends into account (Alresheed et al., 2013; Ma, 2006). Values below 70% suggest a small effect size, values between 70% and 90% a medium effect size, and values above 90% a large effect size.

Results

Expressive Vocabulary

At the beginning of the study, all seven children showed low baseline scores in expressive vocabulary. The average vocabulary size ranged mostly between 0 and 1 word; several children were unable to correctly name any of the vocabulary words at that point. During the first intervention phase (Storytelling, Phase B), performance remained largely unchanged, with means still in the range of 0 to 1 word. Five of the children remained at a score of 0 despite the intervention. Only Aylin and Tugce showed minimal progress: Aylin improved from $M = 1.00$ to $M = 1.25$, and Tugce from $M = 0.50$ to $M = 0.60$.

Significant gains in expressive vocabulary were observed only with the start of the BC phase (storytelling combined with DI using flashcards). All children achieved higher scores in this phase than in the preceding phases A and B (Table 2). The mean values ranged from $M = 2.13$ for Aras to $M = 15.25$ for Aylin. Additionally, all children increased their maximum scores during this phase; the number of correctly named words ranged from 7 to 29 out of a possible 30. Aylin achieved the highest performance level, with a mean of 15.25 and a maximum of 27 correctly named words. Ivan, Defne, and Alara also showed substantial gains, with mean scores of $M = 13.30$, $M = 13.91$, and $M = 12.00$, respectively. Tugce, Leith, and Aras also improved, though their gains were smaller ($M = 4.50$; $M = 5.00$; $M = 2.13$).

The calculated overlap indices reflect these results. When comparing phases A and B, none of the children showed significant changes. The Tau-U values were close to zero or could

not be calculated due to a lack of variance, and the PEM-T reached 50% only for Aylin (p = .69), while for all other children it was 0% (p = 1.00).

In contrast, the comparison between phases B and BC revealed consistently high effect sizes. The Tau-U values showed moderate to strong effects across the board, between 0.73 and 0.97 (p < .001 for all children). The PEM-T supported these findings, with high effect sizes in four out of seven children (100%, p < .01 or p < .001) and moderate effect sizes in three children (87.5%, p < .05; 75%, p = .14).

Table 2

Descriptive Data and Overlap Indices for Expressive Vocabulary of Each Participant

	N_A	N_B	N_BC	M_A (SD)	M_B (SD)	M_BC (SD)	Max A	Max B	Max BC	Tau-U AB (p)	Tau-U BBC (p)	PEM-T AB (p)	PEM-T BBC (p)
Aylin	5	4	8	1.00 (0.00)	1.25 (0.96)	15.25 (8.52)	1	2	27	0.15 (.61)	0.85 (<.001)	50.00% (.69)	100.00% (<.01)
Ivan	3	3	10	0.00 (0.00)	0.00 (0.00)	13.30 (6.85)	0	0	25	N.C. (N.C.)	0.97 (<.001)	0.00% (1.00)	100.00% (<.001)
Defne	4	5	11	0.00 (0.00)	0.00 (0.00)	13.91 (7.42)	0	0	29	N.C. (N.C.)	0.94 (<.001)	0.00% (1.00)	100.00% (<.001)
Alara	5	5	8	0.00 (0.00)	0.00 (0.00)	12.00 (6.37)	0	0	26	N.C. (N.C.)	0.89 (<.001)	0.00% (1.00)	100.00% (<.01)
Tugce	4	5	8	0.50 (0.57)	0.60 (0.55)	4.50 (2.07)	1	1	7	0.15 (.62)	0.73 (<.001)	0.00% (1.00)	87.50% (<.05)
Leith	7	4	8	0.00 (0.00)	0.00 (0.00)	5.00 (3.21)	0	0	11	N.C. (N.C.)	0.75 (<.001)	0.00% (1.00)	87.50% (<.05)
Aras	7	4	8	0.00 (0.00)	0.00 (0.00)	2.13 (2.42)	0	0	7	N.C. (N.C.)	0.82 (<.001)	0.00% (1.00)	75.00% (.14)

*Note.* Measurements exclusive missing values (N); Baseline (A); Storytelling (B); Storytelling + Flashcards (BC); Mean (M); Standard Deviation (SD); Maximum (Max); Percentage of data exceeding a median trend (PEM-T); Not Calculable (N.C.)

The visual analysis of the data confirmed the descriptive findings (Figure 2). During phases A and B, all children showed a predominantly flat progression with stable zero or minimal scores. A clear increase became evident only in the BC phase. For four children (Aylin, Defne, Ivan, Alara), this increase was particularly abrupt, indicating a phase-related intervention effect. Additionally, data variability increased noticeably during this phase, while the performance level remained elevated across multiple measurement points.

Figure 2

Expressive Vocabulary of Each Participant
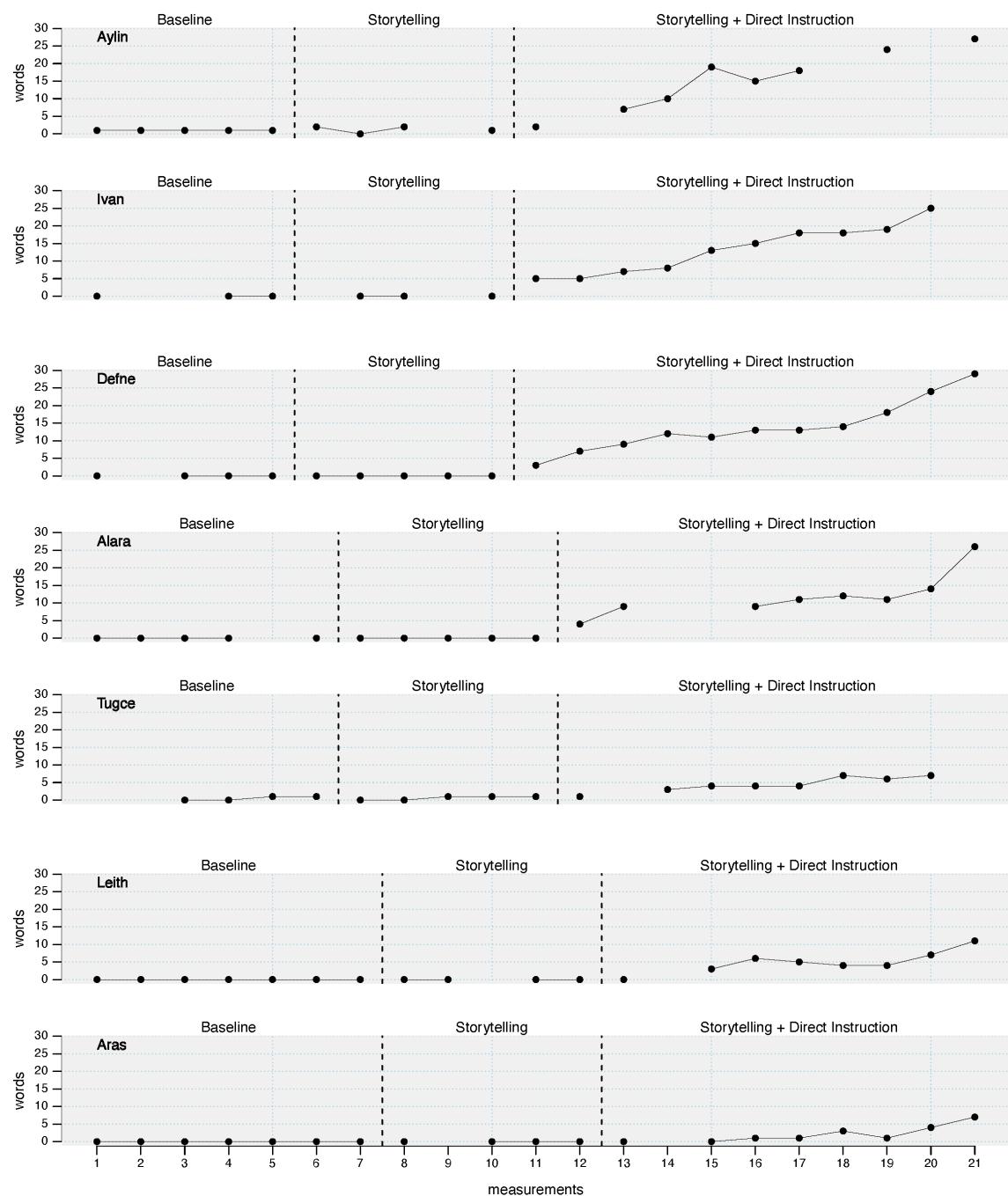
Grapheme-Phoneme Correspondence

Regarding the GPC results, notable differences in initial baseline scores were already evident at the start of the study (Table 3). While Aylin nearly almost reached a ceiling effect in the baseline phase with a maximum of 20 out of 22 possible sounds and a high mean score ($M = 19.20$), Ivan, Defne, Tugce, and Aras demonstrated only very limited knowledge in this area, with mean scores of $M = 2.67$, $M = 1.00$, $M = 0.25$, and $M = 0.00$, respectively. Leith and Alara showed moderate baseline performance with $M = 5.71$ and $M = 9.80$.

As in the case of expressive vocabulary, only minimal changes were observed during the B phase (storytelling) in the area of GPC. As expected, Aylin maintained a high performance level with a mean of $M = 19.25$. Alara, Leith, and Ivan showed slight increases, while Defne and Aras remained unchanged. Tugce even experienced a slight decline, with her mean score dropping from $M = 0.25$ to $M = 0.00$.

With the onset of the BC phase (storytelling + DI), all children improved their GPC scores. Defne increased her mean score to $M = 10.60$ and reached a maximum of 17 sounds. Alara improved to $M = 17.13$ with a maximum of 20 sounds. Ivan and Leith also made gains, achieving maximum scores of 13 and 12 sounds, respectively. Aras scored for the first time in this phase, with a mean of $M = 2.13$. Aylin remained at a very high level with $M = 21.13$, while Tugce made only minor progress, reaching $M = 0.38$.

The overlap indices confirmed these findings. The comparison between Phase A and B showed mostly low to moderate effects, with Tau-U values ranging from -0.20 (Aylin) to 0.54 (Ivan) without significance; for two children, no value could be calculated due to lack of variance. The PEM-T was 0% for most children ($p = 1.00$); only Tugce ($p < .05$) and Ivan ($p = .13$) reached 100%.

In contrast, the comparison between phases B and BC showed a marked increase in all overlap indices. Tau-U values ranged from 0.38 (Aylin) to 0.94 (Defne), with significant effects for all children except Aylin ($p = .11$). The PEM-T indicated low effects for Tugce (37.5%) and Aras (50%), but high effects for Defne, Alara, and Leith (each 100%), all of which were statistically significant.

Table 3

Descriptive Data and Overlap Indices for Grapheme-Phoneme Correspondence of Each Participant

	N_A	N_B	N_BC	M_A (SD)	M_B (SD)	M_BC (SD)	Max A	Max B	Max BC	Tau-U AB (*p*)	Tau-U BBC (*p*)	PEM-T AB (*p*)	PEM-T BBC (*p*)
Aylin	5	4	8	19.20 (0.84)	19.25 (0.96)	21.13 (0.64)	20	20	22	-0.20 (.50)	0.38 (.11)	0.00% (1.00)	75.00% (.14)
Ivan	3	3	10	2.67 (0.58)	3.67 (0.58)	7.40 (3.63)	3	4	13	0.54 (.15)	0.80 (<.001)	100.00% (.13)	80.00% (.05)
Defne	4	5	11	1.00 (0.00)	1.00 (0.00)	10.18 (4.64)	1	1	17	N.C. (N.C.)	0.94 (<.001)	0.00% (1.00)	100.00% (<.001)
Alara	5	5	8	9.80 (1.10)	10.60 (1.52)	17.13 (2.17)	11	12	20	0.26 (.33)	0.70 (<.001)	0.00% (1.00)	100.00% (<.01)
Tugce	4	5	8	0.25 (0.50)	0.00 (0.00)	0.38 (0.52)	1	0	1	-0.12 (.70)	0.62 (<.05)	100.00% (<.05)	37.50% (.86)
Leith	7	4	8	5.71 (0.76)	5.75 (1.71)	8.25 (1.83)	7	8	12	0.02 (.93)	0.51 (<.05)	50.00% (.69)	100.00% (<.01)
Aras	7	4	8	0.00 (0.00)	0.00 (0.00)	2.13 (2.70)	0	0	7	N.C. (N.C.)	0.72 (<.001)	0.00% (1.00)	50.00% (.64)

*Note*. Measurements exclusive missing values (N); Baseline (A); Storytelling (B); Storytelling + Flashcards (BC); Mean (M); Standard Deviation (SD); Maximum (Max); Percentage of data exceeding a median trend (PEM-T); Not Calculable (N.C)

The visual analysis of the data highlighted the heterogeneous picture of GPC (Figure 3). Aylin showed consistently high values across all phases with an almost horizontal trend. For Alara and Leith, who had moderate baseline values, consistent increases were observed across multiple measurement points during the BC phase. For Defne and Ivan, who started with low baseline values, a sudden increase in performance level became apparent with the beginning of the BC phase. Tugce and Aras, despite occasional outliers, remained at an overall low performance level.

Figure 3

Grapheme-Phoneme Correspondence of Each Participant
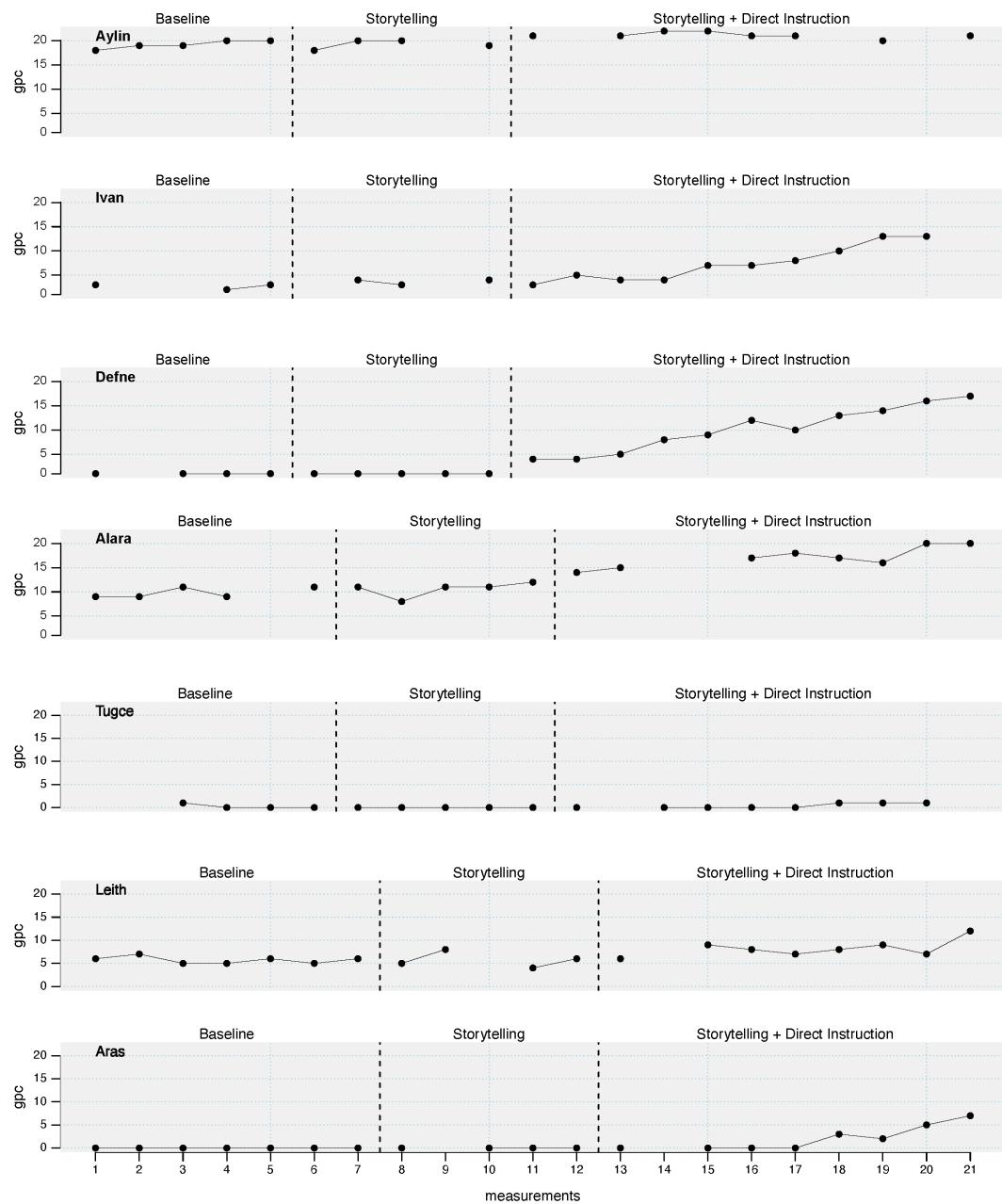
Social Validity

The results on social validity are based on the evaluations of the children and educators. Two children (Ivan and Tugce) were absent at the time of the final measurement, so no data are available for them. The remaining five children rated all items on the questionnaire consistently positively, giving the maximum score (2 points; green thumbs-up).

The educators also evaluated the intervention positively overall. They particularly highlighted the narrative character of the storytelling, which received the highest rating (4 points). Lower ratings were given to the items related to the scope of the intervention and to the perceived changes before and after the intervention (average < 2 points). Higher ratings (2–3 points) were given for items concerning the children's enjoyment of the sessions, the perceived effectiveness, and their own willingness to apply the method.

Discussion

Main Findings

The aim of the present study was to investigate the effectiveness of a storytelling intervention with and without DI on GPC and expressive vocabulary in preschool children learning German as L2. Based on the descriptive and visual analyses, differentiated conclusions can be drawn regarding the effects of the two intervention models.

The key findings show that the storytelling method without DI did not lead to any notable improvements in expressive vocabulary. The children generally remained at their baseline levels, with only minimal gains observed in a few participants. Significant vocabulary gains were only achieved when the storytelling was combined with DI through the use of flashcards. For several children, this effect became apparent immediately at the beginning of the BC phase. Descriptive data showed increases in mean scores of up to 14 words; the overlap indices mostly indicate high effect sizes. These results provide strong indication for the effectiveness of the combined intervention, which is further supported by visual analyses – particularly through abrupt level shifts and sustained elevated performance.

A similar pattern emerged in the area of GPC. During the storytelling phase without DI (Phase B), most children's performance remained largely unchanged. Isolated gains were not statistically significant. Clear improvements were only observed with the combined use of storytelling and DI, particularly among children with very low initial scores. In several cases, mean values in the BC phase were five to ten times higher than in the baseline. These results were supported by both the overlap indices and the visual analysis, and can be clearly attributed to the combined intervention using flashcards.

The results of this study are consistent with previous research findings that emphasize the importance of explicit language support for children learning German as L2 (Dixon et al., 2022; Valcárcel Jiménez et al., 2024). The findings show that effective language promotion requires the integration of semantic, phonological, and orthographic representations of the learning content (Carson et al., 2018; Lesaux et al., 2007; Valcárcel Jiménez et al., 2024). While storytelling, with its narrative context and repeated retelling, serves as a motivating method to support language development (Isbell et al., 2004), it is only through DI that focused attention on the target vocabulary and GPC becomes possible. The use of flashcards provides an important visual complement to the auditory input provided through storytelling. The present findings therefore support the assumption that significant effects on expressive vocabulary and GPC can only be achieved through the combination of both methods (Ehri, 2005).

While storytelling through the narrative reading of stories can serve as a motivational factor in supporting language development, it is DI that places a focused emphasis on target vocabulary and GPC (Gallagher et al., 2019). Without this guided attention to specific linguistic elements, children may process the content only superficially. Furthermore, merely listening can be particularly challenging for children with limited attention spans, potentially leading to reduced motivation and, consequently, diminished focus on the learning content (Lenhart et al., 2020). Neurobiological research supports this perspective, showing that socially interactive reading or guided text experiences enhance language competence by boosting attention and engagement (Horowitz-Kraus et al., 2023). The current findings, therefore, support the conclusion that listening to stories alone is insufficient to capture and sustain children's attention in a motivational and engaging way that facilitates effective language learning. DI is necessary to direct children's focus toward specific linguistic elements to ensure deeper processing and greater learning gains (Gallagher et al., 2019).

Ultimately, it can be observed that a so-called Robin Hood effect occurred (Häfner et al., 2017): children with weaker initial performance benefited the most from the intervention, while those who already showed relatively high performance at the outset made comparatively smaller gains. This effect is particularly positive in the context of special education, especially in the area of learning support or learning difficulties. These children are at risk of experiencing a widening gap in achievement compared to their peers over time (Grünke & Bracht, 2025). Targeted interventions, such as the one implemented in this study, can help counteract existing delays and improve opportunities for participation in the education system (Becker & Lauterbach, 2016; OECD, 2020; Valcárcel Jiménez et al., 2024).

The results of the social validity questionnaires show an overall predominantly positive assessment by both children and educators. All children who completed the questionnaire rated every item with the maximum score ("thumbs up"), indicating high acceptance, enjoyment of the intervention, and a subjective sense of learning progress. The educators also evaluated the intervention positively overall, though certain aspects were assessed more critically. The narrative character of the storytelling was consistently highlighted as especially positive.

The scope of the intervention and the immediate perceptibility of changes were viewed more critically. Some educators found the intervention to be relatively demanding. However, this perception contrasts with the documented learning gains among the children, only two of whom still showed low results at the end. The children's enjoyment of the sessions and their perceived learning progress were also confirmed by the educational staff, which highlights their general willingness to implement the intervention in future practice. Overall, the feedback indicates a high level of acceptance and pedagogical effectiveness of the intervention, even if not all effects are immediately and equally visible in day-to-day preschool settings.

Limitations

Despite the promising results, several limitations must be considered when interpreting the findings. One limitation concerns the heterogeneity and insufficient assessment of the children's language backgrounds. Although all participants were learning German as L2, the extent to which German was spoken in their daily lives, whether German was used as a family language, or whether there were additional language contacts outside of the preschool setting was not systematically recorded. Furthermore, the children's competencies in their respective L1 were not assessed, which would have been important for a more comprehensive understanding of their individual linguistic profiles. In addition, external influencing factors – such as parental involvement, individual developmental trajectories, or educational activities outside the intervention – cannot be fully ruled out.

Another limitation lies in the use of the same vocabulary words across all phases, making sequence effects possible. It is also not possible to clearly attribute the observed effects to storytelling, DI, or the combination of both, due to the combined design of the BC phase.

Finally, an important limitation relates to the sustainability of the effects. Since no follow-up assessments were conducted, no conclusions can be drawn about the long-term impact of the intervention on the children's language development (Engel & Schutt, 2016). While the absence of a long-term follow-up is acknowledged as a limitation, implementing it would be challenging due to the children's diverse school placements after preschool. Nonetheless,

closely monitoring the transition to primary school would be valuable for assessing the lasting impact of early language interventions.

Implications for Future Research and Practice

The results of this study suggest that the use of storytelling alone was significantly less effective – if not largely ineffective – compared to the combined use of storytelling with DI. The markedly stronger effects observed in the combined intervention indicate that the explicit, instructive component may play a crucial role in the observed learning gains.

Future research should therefore specifically investigate to what extent incidental learning through storytelling truly contributes to the acquisition of expressive vocabulary and GPC, or whether the effects are primarily driven by DI. A follow-up study could incorporate an additional C phase that focuses solely on DI – e.g., using flashcards – in order to examine its isolated effectiveness.

For educational practice, the feasibility of the intervention is of particular importance. The combination of storytelling with DI represents a low-threshold, practice-oriented method for implementing targeted language support at an early stage in preschool settings. The use of stories is especially well suited to fostering dialogue between educators and children, as it takes place in a natural and motivating context (Barwasser et al., 2021, 2022; Isbell et al., 2004). Additionally, the time and material effort required to prepare stories and flashcards is comparatively low. The content of the stories can be individualized and adapted to the children's everyday experiences which enhances identification with the story and further boosts motivation (Flynn, 2004).

The present findings demonstrate that early language support in preschool can already be effective for children learning German as L2, and can thus make a significant contribution to school readiness. This leads to the implication that language support measures should be firmly integrated into everyday preschool routines from an early age in order to prevent the manifestation of language difficulties.

The Robin Hood effect observed in this study – meaning greater learning gains among children with lower initial proficiency – highlights the potential of combined instructional approaches such as the integration of storytelling and DI. Such interventions can serve as effective tools to reduce existing language delays among children with German as L2 and to counteract educational disadvantages at an early stage, giving these children a better chance to catch up with their German-speaking peers.

Conclusion

The present single-case study investigated the effectiveness of an early childhood language intervention on phonological awareness and expressive vocabulary in preschool children learning German as L2. The core of the intervention was the storytelling method, supplemented by DI using flashcards. The results clearly show that the combination of narrative instruction and explicit teaching represents an effective approach to supporting L2 acquisition. While incidental learning through storytelling alone did not result in significant performance gains, the combination with intentional learning through targeted instruction led to stable and substantial progress in nearly all participating children.

These findings align with existing research on the interplay of implicit and explicit learning processes in early childhood education and highlight the effectiveness of multimodal instructional approaches at an early age. The results suggest that narrative-based formats should be supplemented with structured and repetitive instruction in order to effectively support core language and pre-literacy skills.

The study demonstrates that targeted language support for children with German as L2 and potential risk of learning difficulties can be both appropriate and effective at the preschool level – that is, before school entry, and before additional educational barriers emerge.

Further research is needed to replicate and refine these findings and to systematically investigate the long-term effects of such interventions.

References

Alresheed, F., Hott, B. L., & Bano, C. (2013). Single subject research: A synthesis of analytic methods. *The Journal of Special Education Apprenticeship*, *2*(1). https://doi.org/10.58729/2167-3454.1015

Anger, C., Betz, J., & Plünnecke, A. (2024). *Bildungsmonitor 2024: Studie im Auftrag der Initiative Neue Soziale Marktwirtschaft (INSM)* [Education Monitor 2024: Study commissioned by the New Social Market Economy Initiative]. INSM.

Baron, L. S., & Arbel, Y. (2022). An implicit-explicit framework for intervention methods in developmental language disorder. *American Journal of Speech-Language Pathology*, *31*(4), 1557–1573. https://doi.org/10.1044/2022_AJSLP-21-00172

Barwasser, A., Bracht, J. & Grünke, M. (2021). A storytelling approach on vocabulary, reading, and letter sound fluency of struggling first graders with German as second language with and without behavioral problems. *Frontiers in Psychology*, *12*, article 683873. https://doi.org/10.3389/fpsyg.2021.683873

Barwasser, A., Skrobek, H., & Grünke, M. (2025). Phonological awareness training on

recoding of at-risk German kindergarten children with and without behavioral difficulties. *International Journal for Research in Learning Disabilities*, *8*(1), 51–65.

Barwasser, A., Urton, K., Knaak, T. & Grünke, M. (2022). Effects of a multicomponent storytellling intervention on the vocabulary acquisition of German L2 adolescents with learning disabilities and peer-relation difficulties. *Insights into Learning Disabilities*, *19*(2), 173–192. https://eric.ed.gov/?id=EJ1357405

Bassetti, B., Mairano, P., Masterson, J., & Cerni, T. (2022). The efficacy of grapheme-phoneme correspondence instruction in reducing the effect of orthographic forms on second language phonology. *Applied Psycholinguistics*, *43*(3), 683–705. https://doi.org/10.1017/S014271642200008X

Becker, R., & Lauterbach, W. (2016). Vom Nutzen vorschulischer Erziehung und Elementarbildung – bessere Bildungschancen für Arbeiter- und Migrantenkinder? [The benefits of pre-school and elementary education - better educational opportunities for working-class and migrant children?] In R. Becker & W. Lauterbach (Eds.), *Bildung als Privileg: Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit* [Education as a privilege: explanations and findings on the causes of educational inequality] (pp. 129–160). Springer VS. https://doi.org/10.1007/978-3-531-91173-1_5

Carson, K. L., Bayetto, A. E., & Roberts, A. F. B. (2018). Effectiveness of preschool-wide teacher-implemented phoneme awareness and letter-sound knowledge instruction on code-based school-entry reading readiness. *Communication Disorders Quarterly*, *41*(1), 42–53. https://doi.org/10.1177/1525740118789061

de Vos, J. F., Schriefers, H., Nivard, M. G., & Lemhöfer, K. (2018), A meta-analysis and meta-regression of incidental second language word learning from spoken input. *Language Learning*, *68*(4), 906–941. https://doi.org/10.1111/lang.12296

Dixon, C., Hessel, A., Smith, N., Nielsen, D., Wesierska, M., & Oxley, E. (2022). Receptive and expressive vocabulary development in children learning English as an additional language: Converging evidence from multiple datasets. *Journal of Child Language*, *50*(3), 610–631. https://doi.org/10.1017/S0305000922000071

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4

Engel, R. J., & Schutt, R. K. (2016). *The practice of research in social work* (4th ed.). SAGE Publications.

Feng, Y., & Webb, S. (2020). Learning vocabulary through reading, listening and viewing. *Studies in Second Language Acquisition*, *42*(3), 499–523.

https://doi.org/10.1017/S0272263119000494

Flynn, S. (2004). Animal stories. In P. Hunt (Ed.), *International companion encyclopedia of children's literature* (2nd ed., pp. 418–435). Routledge. https://doi.org/10.4324/9781315015729-31

Gallagher, M. A., Barber, A. T., Beck, J. S., & Buehl, M. M. (2019). Academic vocabulary: Explicit and incidental instruction for students of diverse language backgrounds. *Reading & Writing Quarterly*, *35*(2), 84–102. https://doi.org/10.1080/10573569.2018.1510796

Grigorakis, I., Kargiotidis, A., Mouzaki, A., & Manolitsis, G. (2022). Early oral language precursors of different types of reading difficulties in a consistent orthography. *Frontiers in Education*, *7*, article 959374. https://doi.org/10.3389/feduc.2022.959374

Grünke, M., & Bracht, J. (2025). Förderung bei Lernstörungen [Support for learning disabilities]. In G. Quenzel, K. Hurrelmann, J. Groß Ophoff & C. Weber (Eds.), *Handbuch Bildungsarmut* [Handbook on educational poverty] (chapter 40–1). Springer VS. https://doi.org/10.1007/978-3-658-44698-7_40-1

Häfner, I., Flunger, B., Dicke, A.-L., Gaspard, H., Brisson, B. M., Nagengast, B., & Trautwein, U. (2017). Robin Hood effects on motivation in math: Family interest moderates the effects of relevance interventions. *Developmental Psychology*, *53*(8), 1522–1539. https://doi.org/10.1037/dev0000337

Horowitz-Kraus, T., Magaliff, L. S., & Schlaggar, B. L. (2023). Neurobiological evidence for the benefit of interactive parent–child storytelling: Supporting early reading exposure policies. *Policy Insights from the Behavioral and Brain Sciences*, *11*(1), 51–58. https://doi.org/10.1177/23727322231217461

Hostetter A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*(2), 297–315. https://doi.org/10.1037/a0022128

Institut zur Qualitätsentwicklung im Bildungswesen [IQB]. (2022). *IQB-Bildungstrends 2021: Kompetenzen und schulische Lernbedingungen im zweiten Pandemie-Jahr* [IQB education trends 2021: Competencies and school learning conditions in the second year of the pandemic]. Waxmann.

Isbell, R., Sobol, J., Lindauer, L., & Lowrance, A. (2004). The effects of storytelling and story reading on the oral language complexity and story comprehension of young children. *Early Childhood Education Journal*, *32*(3), 157–163. https://doi.org/10.1023/B:ECEJ.0000048967.94189.a3

Junttila, K., & Ylinen, S. (2020). Intentional training with speech production supports

children's learning the meanings of foreign words: A comparison of four learning tasks. *Frontiers in Psychology*, *11*, article 1108. https://doi.org/10.3389/fpsyg.2020.01108

Karipidis, I. I., Pleisch, G., Brandeis, D., Roth, A., Röthlisberger, M., Schneebeli, M., Walitza, S., & Brem, S. (2018). Simulating reading acquisition: The link between reading outcome and multimodal brain signatures of letter–speech sound learning in prereaders. *Scientific Reports*, *8*, article 7121. https://doi.org/10.1038/s41598-018-24909-8

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification*, *1*(4), 427–452. https://doi.org/10.1177/014544557714001

Kitzinger, A. (2023). *METACOM Symbole*. https://www.metacom-symbole.de/index.html

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

Law, J., Clegg, J., Rush, R., Roulstone, S., & Peters, T. J. (2019). Association of proximal elements of social disadvantage with children's language development at 2 years: An analysis of data from the Children in Focus (CiF) sample from the ALSPAC birth cohort. *International Journal of Language & Communication Disorders*, *54*(3), 362–376. https://doi.org/10.1111/1460-6984.12442

Ledford, J. R., & Gast, D. L. (Eds.). (2024). *Single case research methodology: Applications in special education and behavioral sciences* (4th ed.). Routledge. https://doi.org/10.4324/9781003294726

Lenhart, J., Lenhard, W., Vaahtoranta, E., & Suggate, S. (2020). More than words: Narrator engagement during storytelling increases children's word learning, story comprehension, and on-task behavior. *Early Childhood Research Quarterly*, *51*, 338–351. https://doi.org/10.1016/j.ecresq.2019.12.009

Lesaux, N. K., Rupp, A. A., & Siegel, L. S. (2007). Growth in reading skills of children from diverse linguistic backgrounds: Findings from a 5-year longitudinal study. *Journal of Educational Psychology*, *99*(4), 821–834. https://doi.org/10.1037/0022-0663.99.4.821

Lewalter, D., Diedrich, J., Goldhammer, F., Köller, O., & Reiss, K. (Eds.). (2023). *PISA 2022: Analyse der Bildungsergebnisse in Deutschland* [PISA 2022: Analysis of educational results in Germany]. Waxmann. https://doi.org/10.31244/9783830998488

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

Mayer, A. (2020). *Test zur Erfassung der phonologischen Bewusstheit und der Benennungsgeschwindigkeit (TEPHOBE): Manual* [Test for the assessment of phonological awareness and rapid naming (TEPHOBE): Manual] (4th ed.). Ernst Reinhardt Verlag. http://www.reinhardt-verlag.de/de/titel/54688/

Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching*, *56*(1), 1–24. https://doi.org/10.1515/iral-2017-0016

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

Nelson, M. C., Cordray, D. S., Hulleman, C. S., Darrow, C. L., & Sommer, E. C. (2012). A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *Journal of Behavioral Health Services & Research*, *39*(4), 374–396. https://doi.org/10.1007/s11414-012-9295-x

Neuman, S. B., Kaefer, T., & Pinkham, A. (2022). Incidental word learning through multiple media: A case for synergy. *Topics in Language Disorders*, *42*(4), 290–303. https://doi.org/10.1097/TLD.0000000000000293

Organisation for Economic Co-operation and Development [OECD] (Ed.). (2023). *PISA 2022 Ergebnisse: Lernstände und Bildungsgerechtigkeit* [PISA 2022 results: Learning levels and educational equity]. Waxmann. https://doi.org/10.3278/6004956w

Organisation for Economic Co-operation and Development [OECD] (Ed.). (2020). *Quality early childhood education and care for children under age 3: Results from the Starting Strong Survey 2018 (TALIS)*. OECD Publishing. https://doi.org/10.1787/99f8bc95-en

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*(3), 255–287. https://doi.org/10.1037/h0084295

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, *42*(2), 284–299. https://doi.org/10.1016/j.beth.2010.08.006

Sachverständigenrat deutscher Stiftungen für Integration und Migration (SVR). (2013). *Hürdenlauf zur Kita: Warum viele Kinder mit Migrationshintergrund von frühkindlicher Bildung ausgeschlossen sind* [Hurdles to kindergarten: Why many children with a migration background are excluded from early childhood education]. SVR GmbH. https://www.svr-migration.de/publikationen/huerdenlauf-zur-kita/

Sammour-Shehadeh, R., Prior, A., & Kahn-Horwitz, J. (2025). Spelling challenges in English

as a foreign language: vowels, digraphs and novel phonemes. *Reading and Writing*. Advance online publication. https://doi.org/10.1007/s11145-025-10659-3

Sanetti, L. M. H., Cook, B. G., & Cook, L. (2021). Treatment fidelity: What it is and why it matters. *Learning Disabilities Research & Practice*, *36*(1), 5–11. https://doi.org/10.1111/ldrp.12238

Senatsverwaltung für Gesundheit, Pflege und Gleichstellung Berlin. (2018). *Grundauswertung der Einschulungsdaten in Berlin 2017* [Basic evaluation of school enrollment data in Berlin 2017]. Referat Gesundheitsberichterstattung, Epidemiologie, Gesundheitsinformationssysteme, Statistikstelle. https://www.berlin.de/sen/gesundheit/service/gesundheitsberichterstattung/

Statistisches Bundesamt (Destatis) (2024a). *Sozialbericht 2024. Ein Datenreport für Deutschland* [Social Report 2024: A data report for Germany]. Bundeszentrale für politische Bildung (bpb). https://www.bpb.de/kurz-knapp/zahlen-und-fakten/sozialbericht-2024/

Statistisches Bundesamt (Destatis). (2024b). *Betreuungsquoten der Kinder unter 6 Jahren (mit und ohne Migrationshintergrund) am 01.03.2023 nach Bundesländern (Mikrozensus-Auswertung)* [Childcare rates for children under 6 (with and without a migration background) on 01.03.2023 by federal state (microcensus evaluation)]. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Soziales/Kindertagesbetreuung/Tabellen/betreuungsquote-migration-unter6jahren-nach-laendern-2023.html

Statistisches Bundesamt (Destatis). (2024c). *Bevölkerung in Privathaushalten 2024 nach Migrationshintergrund (Mikrozensus-Auswertung)* [Population in private households in 2024 by migration background (microcensus evaluation)]. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Tabellen/migrationshintergrund-alter.html

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis or correlation studies. *Language Learning*, *69*(39), 559–599. https://doi.org/10.1111/lang.12343

Valcárcel Jiménez, M., Yumus, M., Schiele, T., Mues, A., & Niklas, F. (2024). Preschool emergent literacy skills as predictors of reading and spelling in Grade 2 and the role of migration background in Germany. *Journal of Experimental Child Psychology*, *244*, article 105927. https://doi.org/10.1016/j.jecp.2024.105927

Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in Psychology*, *3*, article 46. https://doi.org/10.3389/fpsyg.2012.00046

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 of German schoolchildren. *Journal of School Psychology*, *66*, 25–40. https://doi.org/10.1016/j.jsp.2017.11.003

Webber, C., Patel, H., Cunningham, A., Fox, A., Vousden, J., Castles, A., & Shapiro, L. (2024). An experimental comparison of additional training in phoneme awareness, letter-sound knowledge and decoding for struggling beginner readers. *British Journal of Educational Psychology*, *94*(1), 282–305. https://doi.org/10.1111/bjep.12641

Wilbert J, & Lüke T (2025). *Scan: Single-case data analyses for single and multiple baseline designs*. https://CRAN.R-project.org/package=scan

Wolery, M., Dunlap, G., & Ledford, J. R. (2011). Single-case experimental methods: Suggestions for reporting. *Journal of Early Intervention*, *33*(2), 103–109. https://doi.org/10.1177/1053815111418235

Yang, N., Shi, J., Lu, J., & Huang, Y. (2021). Language development in early childhood: Quality of teacher-child interaction and children's receptive vocabulary competency. *Frontiers in Psychology*, *12*, article 649680. https://doi.org/10.3389/fpsyg.2021.649680

Appendix F Declaration of Own Performance

Article 1	
Hertel, S., Bracht, J., Calhoon, M. B., Grünke, M., & Barwasser, A. (2024). Effects of an adapted peer-assisted learning strategies reading programme on reading fluency and reading comprehension of secondary students with or at-risk for reading disabilities. *European Journal of Special Needs Education*, 1–17. https://doi.org/10.1080/08856257.2024.2402166	
Research idea	with co-authors
Current state of research	co-authors
Design/Measurements	co-authors
Study Overall control	with co-authors
Data analysis	co-authors
Publication	with co-authors

Article 2	
Bracht, J., Hoff, S., Grünke, M., & Barwasser, A. (2025_accepted). Enhancing reading competencies of German as a second language learners through an adapted peer-assisted learning strategies reading programme. *Insights into Learning Disabilities*.	
Research idea	with co-authors
Current state of research	overall control
Design/Measurements	overall control
Study Overall control	overall control
Data analysis	with co-authors
Publication	with co-authors

Article 3

Barwasser, A., Bracht, J., Lenz, B., Gürçay, I., Hoff, S., & Grünke, M. (2025) Effects a peer-tutorial story-map intervention on the reading and writing of students with and without German as a second language. *Reading & Writing*, *38*, 1337–1357. https://doi.org/10.1007/s11145-024-10565-0

Research idea	co-authors
Current state of research	with co-authors
Design/Measurements	co-authors
Study Overall control	with co-authors
Data analysis	co-authors
Publication	with co-authors

Article 4

Barwasser, A., Bracht, J., & Grünke, M. (2021). A storytelling approach on vocabulary, reading and letter sound fluency of struggling first graders with German as second language with and without behavioral problems. *Frontiers in Psychology*, 12, article 683873. https://doi.org/10.3389/fpsyg.2021.683873

Research idea	with co-authors
Current state of research	co-authors
Design/Measurements	co-authors
Study Overall control	co-authors
Data analysis	with co-authors
Publication	with co-authors

Article 5

Bracht, J., Wasko, L., Grünke, M., & Barwasser, A. (2025_accepted). Storytelling with and without direct instruction on grapheme-phoneme correspondence and vocabulary of pre-school children learning German as a second language. *Learning Disabilities: A Contemporary Journal*.

Research idea	overall control
Current state of research	overall control
Design/Measurements	with co-authors
Study Overall control	overall control
Data analysis	with co-authors
Publication	with co-authors

Appendix G Declaration of Independence (according to §11 (1) 8)

Humanwissenschaftliche Fakultät
Universität zu Köln

Eidesstattliche Erklärung nach § 11 (1) 8 der Promotionsordnung der Humanwissenschaftlichen Fakultät vom 18.12.2018
(Die Erklärung ist Bestandteil des Antrags auf Eröffnung des Promotionsverfahrens)

☒ **für eine monographische Dissertation mit Teilpublikationen *oder***
für eine kumulative Dissertation:

Ich, Janine Ilona Bracht, geboren am ▮▮▮▮▮▮▮, versichere eidesstattlich, dass ich die von mir vorgelegte Dissertation selbstständig und ohne unzulässige Hilfe angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit einschließlich Tabellen, Karten und Abbildungen, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe sowie dass diese Dissertation noch keinem anderen Fachbereich zur Prüfung vorgelegen hat. Die Promotionsordnung ist mir bekannt. Die von mir vorgelegte Dissertation ist von Professor Dr. Matthias Grünke betreut worden.

▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮

(Ort, Datum)   (Unterschrift)

Appendix H Declaration (according to §11 (1) 7)

Humanwissenschaftliche Fakultät
Universität zu Köln

Erklärung nach § 11 (1) 7 der Promotionsordnung der Humanwissenschaftlichen Fakultät vom 18.12.2018
(Die Erklärung ist Bestandteil des Antrags auf Eröffnung des Promotionsverfahrens)

Hiermit erkläre ich, Janine Ilona Bracht, geboren am ████████, dass ich bisher keinen erfolgreichen oder erfolglosen Versuch zum Erwerb des Doktorgrades an der Humanwissenschaftlichen Fakultät, einer anderen Fakultät, oder einer anderen Hochschule unternommen habe. Ebenso befinde ich mich diesbezüglich nicht in einem schwebenden Verfahren.

(Ort, Datum)   (Unterschrift)

Appendix I List of Publications

*Barwasser, A., **Bracht, J.**, & Grünke, M. (2021). A storytelling approach on vocabulary, reading and letter sound fluency of struggling first graders with German as second language with and without behavioral problems. *Frontiers in Psychology*, *12*, article 683873. https://doi.org/10.3389/fpsyg.2021.683873

*Barwasser, A., **Bracht, J.**, Lenz, B., Gürçay, I., Hoff, S., & Grünke, M. (2025) Effects a peer-tutorial story-map intervention on the reading and writing of students with and without German as a second language. *Reading & Writing*, *38*, 1337–1357. https://doi.org/10.1007/s11145-024-10565-0

***Bracht, J.**, Hoff, S., Grünke, M., & Barwasser, A. (2025_accepted). Enhancing reading competencies of German as a second language learners through an adapted peer-assisted learning strategies reading programme. *Insights into Learning Disabilities*.

***Bracht, J.**, Wasko, L., Grünke, M., & Barwasser, A. (2025_accepted). Storytelling with and without direct instruction on grapheme-phoneme correspondence and vocabulary of preschool children learning German as a second language. *Learning Disabilities: A Contemporary Journal*.

*Frangenberg, Y., Grünke, M., **Bracht, J.**, Jochims, A., Barwasser, A., & Hord, C. (2023). Assessing the Impact of International Scholarly Journals on Learning Disabilities: An Analysis Using a Google-Based Journal Impact Factor. *Insights into Learning Disabilities*, *20*(2), 123–132.

*Grünke, M., & **Bracht, J.** (2025). Förderungen von Lernstörungen. In G. Quenzel, K. Hurrelmann, J. Groß Ophoff & C. Weber (Eds.), *Handbuch Bildungsarmut* (living reference work). Springer VS. https://doi.org/10.1007/978-3-658-44698-7_40-1

*Grünke, M., Gürcay, I., **Bracht, J.**, Jochims, A., Schulden, M., Barwasser, A., & Duchaine, E. (2024). Enhancing grapheme-phoneme correspondence learning: A single-case study using picture mnemonics. *International Electronic Journal of Elementary Education*, *16*(3), 417–425. https://doi.org/10.26822/iejee.2024.342

*Grünke, M., Klöpfer, C., Wellmann, T., & **Bracht, J.** (2020). "I Went on the Mound and I Felt Like I Didn't Have a Learning Disability": A qualitative study of a former professional baseball player with learning disabilities. *Insights into Learning Disabilities*, *17*(2), 109–128.

*Grünke, M., Nobel, K., & **Bracht, J.** (2019). Effects of the STOP and LIST strategy on the writing performance of struggling fourth graders. *Journal of Education and Learning*, *8*(2), 1–13. https://doi.org/10.5539/jel.v8n2p1

*Hertel, S., **Bracht, J.**, Calhoon, M. B., Grünke, M., & Barwasser, A. (2024). Effects of an adapted peer-assisted learning strategies reading programme on reading fluency and reading comprehension of secondary students with or at-risk for reading disabilities. *European Journal of Special Needs Education*, 1–17. https://doi.org/10.1080/08856257.2024.2402166

*Zöller, E., Klöpfer, C., **Bracht, J.**, Müller, P., & Grünke, M. (2025). Swinging for Success: A Qualitative Study on Baseball's Role in Supporting a Player With ADHD. *Disabilities*, *5*(1), article 2. https://doi.org/10.3390/disabilities5010002

---

peer-reviewed