frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 04 November 2024
DOI 10.3389/feduc.2024.1378493

Using an app-based screening tool to predict deficits in written word spelling at school entry

Martin Schöfl[1,2]*, Gabriele Steinmair[1], Sabine Zepnik[1] and Christoph Weber[1,2]

[1]Department of Educational Sciences, University of Education Upper Austria, Linz, Austria, [2]Research Institute for Developmental Medicine, Johannes Kepler University Linz, Linz, Austria

Check for updates

OPEN ACCESS

EDITED BY
Adrian Pasquarella,
University of Delaware, United States

REVIEWED BY
Diana Alves,
University of Porto, Portugal
Bita Moradi,
University of Delaware, United States

*CORRESPONDENCE
Martin Schöfl
martin.schoefl@jku.at

RECEIVED 29 January 2024
ACCEPTED 30 September 2024
PUBLISHED 04 November 2024

CITATION
Schöfl M, Steinmair G, Zepnik S and Weber C (2024) Using an app-based screening tool to predict deficits in written word spelling at school entry. *Front. Educ.* 9:1378493.
doi: 10.3389/feduc.2024.1378493

Introduction: The first year of schooling is crucial for the further development of spelling abilities in children, which makes early assessment and intervention essential. The aim of this study was to develop and validate an efficient and cost-free screening tool for identifying spelling problems in community school settings around the time of school entry.

Methods: A broad range of precursors of spelling (vocabulary, grammar, letter knowledge, phonological awareness, phonological working memory, rapid automatized naming) were assessed in 522 Austrian first graders (6–7 years of age) in the first weeks of schooling. At the end of first grade, spelling abilities were assessed by newly developed spelling tasks based on the trochaic foot. By applying logistic regression with the least absolute shrinkage and selection operator (LASSO), we aimed to select a set of important predictors of spelling problems at the end of grade 1 (i.e., scoring below the 16th percentile in the spelling test).

Results: Our analysis identified letter knowledge (i.e., an aspect of phonological information processing) and sentence repetition (i.e., a measure of grammatical knowledge) as important predictors of spelling problems. The screening tool has acceptable diagnostic accuracy [area under the curve (AUC) = 0.0.725 and DeLong 95% CI (0.666, 0.784)]. Further analyses indicated that the AUC differs neither between boys and girls nor between children with and without German as their first language.

Discussion: These results suggest that administering the screening tool during the first weeks of schooling is a valid approach to identifying spelling deficits, which in turn enables early targeted pedagogical interventions. Practical implications for spelling instructions are discussed.

KEYWORDS

Schnapp spelling test, multilingual first graders, phonological processing, language abilities, graphematic trochaic foot

Introduction

The term *spelling* is used ambiguously: on the one hand, spelling is often associated with the term *oral spelling*, which means memorizing a word and naming its individual letters (Treiman and Bourassa, 2000). On the other hand, it is used to describe the conversion of spoken words by young children into phonemes and thence into a chain of letters (Treiman et al., 2019). It is not clear whether the latter is also related to knowledge about an orthographic form (Treiman et al., 2019). Since spelling generally deals with correctness (Treiman et al., 2023), which also matches our approach, we use the term spelling when we refer to *writing* *down words correctly*, and include implicit and systematic knowledge of the graphematic word structure in German (Bredel, 2015a; Primus, 2010) as a basic principle.

Spelling prediction relies predominantly on phonological information processing (PI) and language comprehension. Embracing a comprehensive approach, we adopt Wagner and Torgesen's (1987) extended definition, which includes various components for word inscription. Phonological information processing, as defined by Wagner and Torgesen (1987), p. 192, entails the utilization of phonological cues—that is, the auditory aspect of language—in both written and oral language processing. This includes sublexical processes, such as letter knowledge (LK), rapid automatized naming (RAN), phonological awareness (PA), and phonological working memory (PWM) (Niolaki et al., 2020; González-Valenzuela et al., 2023). The related research literature (for an overview see Appendix 1) considers as relevant not only precursors related to phonological information processing, but also lexical precursors, such as visual short-term memory (VSTM), visual attention span (VAS) (Niolaki et al., 2020; Niolaki et al., 2024), oral language skills, including vocabulary (V) and grammar (GR) (Kim et al., 2013; von Goldammer et al., 2010; Fricke et al., 2016), and early oral spelling skills (ESS) (Treiman et al., 2023), which are considered foundational for orthographic spelling. We discern that, within English-language research, phonological awareness emerges as a key predictor. For other writing systems, however, the significance of RAN and oral language proficiency is increasingly acknowledged.

Learning to spell is a slow process; correct spelling usually requires effort, regardless of the orthography in question, and some children need extra support (Nag et al., 2019). Because spelling continues to be a highly valued skill (Pan et al., 2021), it is important to find sensitive, accurate methods for assessing children's spellings (Treiman and Kessler, 2004) or—even earlier—precursors for word-writing development and underlying cognitive processes that children use when they write (Caravolas et al., 2001; Kessler et al., 2013; Treiman et al., 2016). Since early orthographic knowledge predicts later orthographic skills (Zarić et al., 2020), early support is essential. Diagnostic and interventional approaches are normally based on theoretical models. Written language acquisition is usually described using so-called phase models (Frith, 1986; for German, e.g., Thomé, 2006; Scheerer-Neumann, 2006). These rely strongly on phoneme-grapheme relations, and require children to listen carefully and write down exactly what they have heard before orthographic writing can be developed. Our approach, however, builds on the alternative graphematical theory introduced by Eisenberg (1989) and Maas (1992) in which the phonological suprasegmental unit—the syllable—is given preference to single letter-sound relations within the written language. This view has been established within graphematics and expanded to the graphematic foot, a sequence of two syllables, a strong and a weak or reduced syllable (Primus, 2010; Evertz, 2016; Maas, 2022; Eisenberg, 2020; Fuhrhop and Peters, 2023). A didactical implementation of this graphematical theory suggests that children are introduced to the graphematic structure based on the syllable and the trochaic foot, and from the beginning of grade 1, they are thus enabled to discover patterns and regularities (Röber, 2013; Bredel, 2016; Bredel et al., 2017).

Existing screening tools

Only a few screening tools exist for use with German orthography: The "LRS-Screening" (Endlich et al., 2019) used in the year preceding formal schooling employs 14 subtests to forecast word-spelling deficits by the end of first grade. Notably, it has a sensitivity of 0.74, a specificity of 0.68, and a positive predictive value (PPV) of 0.27. The "LRS-Screening" is administered individually in a traditional paper-and-pencil format and takes approximately 30 min plus additional scoring time. It lacks an engaging cover story, which potentially limits its appeal. Another screening tool commonly used prior to school entry is the "Bielefelder Screening" (Jansen et al., 1999). Although its manual reports promising predictive values, independent studies have failed to replicate these findings. The Phonological Awareness-Reading and Spelling Screening (PB-LRS; 80) is designed for group assessment during the final year of kindergarten. It boasts a sensitivity of 63%, a specificity of 87%, and a PPV of 36%, but it requires approximately 60 min to administer. Another established tool, the "A Tour of Hörhausen" (Martschinke et al., 2001), comprises phonological assessments individually at the beginning and midpoint of first grade. Prognostic validity studies based on a sample of 375 children revealed a specificity exceeding 80%, with sensitivity ranging from 38 to 48%. Assessment takes approximately 40 min. The only app-based screening is a tool called "Förderorientieres Schuleingangsscreening" (BMBWF, 2024), which is supported by the Austrian government and consists of tasks that concern phonological awareness and letter knowledge (Jöbstl et al., 2022). This forms part of a broader battery of seven modules that also address precursors of numeracy development, graphomotor skills and executive functions. This assessment is of acceptable validity and presented in a motivating format, but lacks important precursors for written language, such as vocabulary and grammar.

In summary, existing screening tools for predicting word spelling difficulties in German are time-consuming to administer or exhibit limited predictive efficacy, as detailed in comprehensive reviews (Marx and Lenhard, 2010). Language as a main factor is, in many cases, not one of the components screened.

The present study

The objective of the present study was to develop and validate an efficient school-entry screening tool for identifying spelling problems at the end of grade one in community school settings. In order to enable broad implementation, the tool had to be administrable by teachers and feasible in terms of interpretation and integration into everyday school life. Therefore, to facilitate assessment and interpretation of results, and following the general trend in psychological and educational assessment towards increasing integration of digital media in practice (Schaumburg, 2015), we developed an app-based screening approach, which in Austria can also be administered in the final year of kindergarten (Thompson, 2015). Unlike existing screening tools, this study considered a broad set of precursors assessed at school entry phonological information processing (initial phoneme detection, letter knowledge, and phonological working memory) and language (vocabulary and grammar) as predictors of spelling. Focusing on the validity of the selected set of predictors, this study evaluated diagnostic accuracy—broadly defined as the ability of the screening to discriminate between children with and without spelling problems. To investigate the generalizability of the screening, an examination was undertaken of the potential for variation in diagnostic accuracy according to sociodemographic variables. These include the child's gender and first language, both of which are known to be associated with spelling as well as spelling precursors. The extant research demonstrates that children from low socioeconomic status (SES) backgrounds exhibit lower verbal academic performance than their peers from high-SES families (Sirin, 2005; White, 1982). Similarly, there is substantial evidence indicating that children from low-SES environments demonstrate weaker orthographic skills, particularly in spelling, in comparison to their higher-SES counterparts (Breit et al., 2016; Niemietz et al., 2023 for German-speaking countries). Large-scale international assessments such as PISA (OECD, 2023) indicate that students who have acquired the language of instruction as a second language, frequently due to a migration background, frequently demonstrate suboptimal performance in spelling. Specifically, research findings suggest that both second-language German learners (Lenhart et al., 2019) and students with a migration background (Henschel et al., 2023) exhibit comparatively weaker spelling abilities than their peers.

In summary, the following research questions guided this study:

> Research question 1: Which predictors assessed at school entry can identify spelling problems at the end of grade 1?
>
> Research question 2: What is the diagnostic accuracy of the selected set of predictors?
>
> Research question 3: Does the diagnostic accuracy of the selected set of predictors vary between sociodemographic groups?

Methods and materials

Sample and recruitment

This prospective study followed children from the beginning to the end of first grade and is part of a longitudinal and ongoing project called SCHNAPP (Detection of risks factors for reading and writing development, see Schöfl et al., 2022). The sample was recruited from various community-based schools in the district of the Upper Austrian capital. The principals of 14 schools were telephoned and then visited in person, and invited to participate in the study project. All of them agreed to participate and informed their teachers of first grade classes ($n = 32$). Parents of 557children (96.7%) provided permission for their child to participate in the study.

Precursors of spelling—phonological information processing and language—were assessed in the first weeks of first grade, before the formal teaching of reading and spelling had begun. In detail, individualized screening of precursors started in autumn 2022 within the first 2 weeks of the school year. Within 3 weeks, 85% of the sample had been assessed. In the two subsequent weeks, children who had been ill or unable to attend within the first testing period were surveyed. Spelling was assessed by a test administered in a classroom setting at the end of first grade. Twenty children were excluded from the analysis because they were listed as (preschoolers[1]) in their first year of learning and did not advance to the first grade at the end of the school year. Eight children were diagnosed during the school year as having special educational needs (SEN). The spelling test was not administered to these children as they had not yet acquired the requisite knowledge of the alphabet. A total of 7 children who were part of the target population for the screening could not be tested (4 children changed school and 3 children were ill, at the first and second test date). Thus, our study used data from $n = 522$ children (=557 (with parental consent) − 20 (preschoolers) − 8 (SEN students) − 7 (missing tests)).[2] For these children all variables used in the current study were available. Figure 1 shows the recruitment pathways and timeline.

The distribution of the children in the final study sample was heterogeneous. The proportion of girls in the sample was 46.7%, which equals the proportion of girls among Upper Austrian first graders in 2022 [$\chi^2(1) = 0.00$, $p > 0.05$; Statistik Austria, 2024]. 73.2% of the children assessed spoke German exclusively as their first language, which also corresponds to the population value (i.e., Upper Austrian first graders) of 73.1% [$\chi^2(1) = 0.01$, $p > 0.05$; Statistik Austria, 2024]. The sample had parents from all educational backgrounds. In about four out of 10 families (42.7%) at least one parent held a university degree. In 21.3% of the families, the highest educational degree was a university entrance qualification. About a third (30.1%) reported a vocational education or training, and in 5.9% of the families neither parent had any qualification beyond compulsory schooling. Note that, comparison of the proportion of parents with the highest educational level in our sample with that of the Upper Austrian parent population of 4th graders in the school year 2017/18[3] shows that families in which at least one parent has a university degree are overrepresented (42.7% vs. 26.5%) and parents with vocational education or training are underrepresented [30.1% vs. 45.8%; $\chi^2(3) = 77.50$, $p < 0.001$]. Thus, our study sample was representative in terms of gender and first language, but not in terms of parental education, which is probably due to the sample having been recruited from many schools in urban areas, where more parents with higher education live (see also Elliott, 2018, p. 245; Yulianti et al., 2023).

Procedure

Before summer vacation, teachers were informed by the research team about the procedure and received an information letter for the parents or legal guardians, including consent forms and questions about parents' educational background, children's first language, and

1 These children were not taught according to the first-grade elementary school curriculum, which means that the children did not learn to write within the study period.

2 Notably, from a missing data perspective (Enders, 2010), the seven cases that were excluded from the study (children who were ill and children who changed school) could be regarded as unit non-response, making an adequate treatment of missing data necessary. However, given the small number of cases with missing values (1.3%) and analyses indicating that missing cases did not significantly differ in study variables from the other cases (p-values >0.088), we decided to exclude these cases (i.e., listwise deletion).

3 We thank the IQS (Federal Institute for Quality Assurance in the Austrian School System) for providing detailed population data.

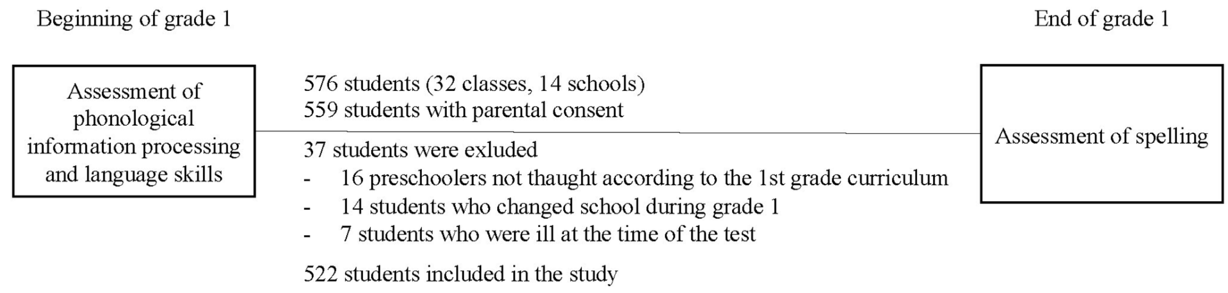

FIGURE 1
Recruitment pathways and timeline.

language use. Teachers entered the IDs on a digital platform for use on a tablet. The investigators, all of whom were student teachers, were enrolled in a student seminar where they learned about the tasks and the testing procedure and for which they received academic credits (towards their study program). Phonological information processing and language skills at the beginning of grade 1 were assessed one on one, with child and instructor seated opposite each other across a table. After a brief welcome, the child was handed a tablet running an app that introduced SCHWUPP the friendly dragon. App navigation was designed such that the child could use it independently. If needed, investigators could help, for instance, to move SCHWUPP across the screen. All instructions essential for the child were recorded as audio files, opened automatically, and repeated if necessary. A yellow background in the app signaled to the test administrator that the child was making test selections independently (e.g., in the phonological tasks). A gray background meant that the test administrator should take the tablet to read and then give instructions. The assessment, including all subtests, took an average of 15.4 min (SD = 4.3) per child. Pauses were possible but rarely needed. The spelling test is also tablet-based (see below). It was administered by the research team and student teachers in a classroom setting.

Measures

Phonological information processing

Letter knowledge

All 26 letters of the alphabet were presented as capital letters on screen in random order. Each page contained three to four letters, and each child was prompted, "I know you have not yet learned these letters at school. Maybe you already know one? If you do, please name it!" The child named letters and tapped on those that they knew. In the case of an error (wrong letter as chosen known), the investigator could correct the error by tapping again. Positive scores were given for letter names or sounds and ticked off on the tablet. Reliability (Cronbach's alpha) of this newly designed task was excellent at 0.96.

Initial phoneme detection

The task was introduced with three practice items, including feedback, followed by 10 test items. Tasks were constructed using high-frequency words from the childLex database (Schroeder et al., 2015) for the youngest age group (6–8 years). We presented each letter visually and as a speech sound simultaneously. The child then had to select from three pictures that which illustrated a word that started with the same first phoneme as the letter given ("Which word begins with I like Ines: Hase, Igel, Spiegel?") (Ines = German name "Ines," Hase = hare, Igel = hedgehog, Spiegel = mirror). Although internal consistency was relatively low at 0.66, this test has been shown to be an important predictor of reading in prior research (Schöfl et al., 2022). Therefore, we also used this subtest in the present study.

Phonological working memory

Phonological working memory was assessed by means of two subtests from a broad-range intelligence test battery [IDS-2; Intelligence and Development Scales for Children and Adolescents (Grob et al., 2009)] which tests the memory of letter-number sequences forward (letter-number span forward) and backward (letter-number span backward). Each child was asked to repeat a series of mixed digits and letters (3-A, 5-M-2) from beginning to end (forward condition) or vice versa (backward condition). The investigator tapped the correct solutions on the tablet. The difficulty level of the tasks was determined by the increasing length of the sequences, and the test was terminated after three unsolved or incorrectly solved tasks. Reliability of this test has been described as excellent; Cronbach's alpha was 0.89 (end of first grade). Retest reliability was rtt = 0.93 (first grade).

Rapid automated naming

To assess RAN, two different stimuli (objects and digits) were used. Analysis of our own data has recently shown effects of both measures assessed at the very beginning of schooling on reading assessed at the end of grade 1 (Schöfl et al., 2023a). The object condition was designed using five high-frequency monosyllabic words (cow, hand, ice, tree, and mouse). First, the items of the RAN tasks were presented visually and by playing an audio file via the app, and the task was repeated. After a training session, the test began: 30 items were presented on the screen in random order over six lines. The investigator pressed a button on the tablet to time the test. When the child reached the last item, the investigator stopped the clock, and the time distance was calculated and stored automatically. The RAN test in the digit condition was based on the work of Denckla and Rudel (1974) following Landerl et al. (2013) and was presented and rated analogously to the object condition with monosyllabic digits (2, 8, 1, 6, 3).

Language

Grammar

Morphosyntactic skills were assessed using an adapted sentence repetition task. The German version was constructed according to the Language Impairment Testing in Multilingual Children (LITMUS) principles by Ibrahim et al. (2018), following the COST Action IS0804. A block of 15 items representing morphosyntactic constructions with varying degrees of complexity was selected and scored according to whether the sentence was completed correctly. Internal consistency was high at 0.857.

Receptive vocabulary

Receptive vocabulary was assessed using the digital form of the Graz Vocabulary Test (GraWo; Seifert et al., 2019), which consists of 30 matching tasks. The child was asked to choose from four pictures the one that best matched the audio-presented word. Reliability data are given for the paper form of the GraWo (Cronbach's Alpha for end of first grade was 0.89; Retest reliability was rtt = 0.93).

Standardized assessment of spelling at the end of the first year

To identify students with spelling deficits, the newly developed and validated SCHNAPP spelling test (Schöfl et al., 2023b) was used. Within a magical cover story about SCHWUPP the dragon, 22 words must be written according to dictation. A stylus pen is used for input to the tablet. The target words and corresponding sentences are played using headphones. The audio files can be repeated upon an input command from the child. The next item is presented only when the child gives an input command, which allows the test to progress at the child's own pace. The vocabulary is based on the childLex corpus, a database of written language for children (Schroeder et al., 2015). In accordance with Schroeder et al. (2015), the words were divided into three groups according to frequency of use: >100 high-frequency, 10–100 medium-frequency, and 1–10 low-frequency. In relation to the writing system, we used a literature-derived hierarchy of vocabulary based on the canonical trochaic form typical of German (see Schöfl et al., 2023b). Schöfl et al. (2023b) showed that the test items conformed to the one-parameter Rasch model, and the test exhibited good reliability (0.86). Further, the test differentiates well in the low ability range, which makes it suitable for identifying spelling deficits. For the current study, test scores one standard deviation below the mean were considered as spelling deficits.

Sociodemographic measures

Mono- vs. bilingualism classification was based on a questionnaire completed by the parents. If they indicated that the first language was German only or that contact with German occurred from birth up to and including the age of 2, then children were classified as monolinguals. Children who had had contact with German only after the age of 2 years were classified as bilingual.

Analyses

First, we performed bivariate analyses. We report point-biserial correlations r_{pb} between the predictors and spelling as well as area under the curve (AUC) values based on receiver operator characteristic (ROC) curve analyses to evaluate the diagnostic accuracy of each predictor. AUCs are global measures of diagnostic accuracy. Values ≥0.9 are regarded as excellent, AUCs ≥0.8 and <0.9 as good, AUCs ≥0.7 and <0.8 as fair, and tests with AUCs <0.7 as poor (Swets, 1988). The pROC package (Robin et al., 2011) in R was used for ROC curve analysis.

Second, we applied multivariate method to simultaneously evaluate the predictive power of spelling precursors. We performed a logistic regression with adaptive variable selection to identify important predictors for inclusion in the screening. Specifically, we used the least absolute shrinkage and selection operator (LASSO; McNeish, 2015; Friedman et al., 2010; for an application of LASSO for the selection of screening variables, see Schöfl et al., 2022; Silverman et al., 2021) implemented in the glmnet R Package (Friedman et al., 2021). LASSO is a statistical learning procedure that serves two purposes: selection of important predictors and attenuation of overfitting. Notably, standard variable selection procedures such as backward and forward selection tend to increase overfitting, which limits the generalizability of the regression results and thus also of a screening algorithm to be developed because sample regression estimates are a blend of signal and noise; sample estimates are therefore greater than population estimates. To counter inflation of the sample estimates and improve generalizability of the results, LASSO applies a penalty term (λ) to the likelihood function. As in backward or forward selection procedures, unimportant predictors are excluded from the regression model. Note that LASSO does not allow any conclusions to be drawn about the statistical significance of a predictor. Although methods for calculating p-values for linear LASSO models have been developed, no such methods are available for logistic models. The predictors selected must be considered important irrespective of whether their regression estimates are statistically significant (McNeish, 2015; Silverman et al., 2021). We z-scored the predictors to make their effects comparable. Thus, the estimates are standardized. As noted, LASSO uses a penalty term (λ) to counter overfitting and zero-out predictors. Following the suggestions in the literature (James et al., 2013), we used 10-fold cross-validation to select the penalty term. Specifically, we chose the λ-value that resulted in the largest AUC (λ_{AUC_max}) to ensure the highest diagnostic accuracy of the prediction model. However, since the maximal AUC may in some cases insufficiently attenuate overfitting, we additionally report the results for a second λ value. We applied the one standard error rule (λ_{AUC_1SE}; selecting the highest value of λ at which the AUC is within one standard error of the greatest AUC; see, e.g., James et al., 2013, p. 214), which selects a model that has a comparable AUC to the model with the highest AUC, but which is more parsimonious (i.e., has a smaller number of predictors). Tenfold cross-validation randomly splits the data into 10 groups (data sets), where one set is used for validation and the remaining nine sets are used for training in the statistical learning procedure. Given this randomness, the results vary with each re-estimation of the model (James et al., 2013). To make our results reproducible and evaluate their robustness, we set five different random seeds and averaged the resulting estimates. For comparison we also report the results of a classical logistic regression using all predictors simultaneously. Finally, we used the estimates of the LASSO models to calculate the probability of spelling deficits at the end of grade 1 (i.e., spelling scores one SD below the mean). These scores, ranging from 0 to 1, were used as screening measures in subsequent analyses.

Third, by applying a bootstrapped test for the comparison of AUC values of paired ROC curves, we compared the diagnostic accuracies of the screening scores λ_{AUC_max} and λ_{AUC_1SE} (Robin et al., 2011).

Fourth, we evaluated whether the ROC curves differed (i) between girls and boys and (ii) between children with and without German as their first language. Differences in ROC curves and AUC values indicate variations in diagnostic accuracy and therefore limit the generalizability of screening (Youngstrom, 2014). We used a bootstrapped test for unpaired ROC curves to compare AUC values between two groups. Further, we applied the Venkatraman (2000) permutation test implemented in the pROC R-package (Robin et al., 2011), which compares actual ROC curves. Note that two ROC curves that do not differ in screening scores have the same sensitivity and specificity across groups, and thus no group-specific cut-off values are required.

Fifth, we used the R-OptimalCutpoints package (López-Ratón et al., 2014) to evaluate several cut-offs that can be used in practice. We report the following diagnostic accuracy statistics: sensitivity (Se; rate of children with spelling deficits identified correctly by the screening process), specificity (Sp; rate of children without spelling deficits identified correctly by the screening), positive predictive value (PPV; rate of screen-positives that show spelling deficits), negative predictive values (NPV; rate of screen-negatives that do not show spelling deficits), and diagnostic likelihood ratios for positive and negative screening results (DLR+ and DLR−, respectively). Unlike predictive values, DLR+ and DLR− are not associated with the prevalence of the problem (here: spelling deficits) under investigation (Pepe, 2004). DLR+ indicates the multiplicative change in the pre-screening odds of scoring one SD below the mean in the spelling test given a positive screening result (i.e., post-screening odds = DLR+ × pre-screening odds). DLR− is the change in the pre-screening odds of scoring one SD below the mean given a negative screening result (post-screening odds = DLR− × pre-screening odds). In the medical literature, DLR+ values ≥10 and DLR− ≤0.1 indicate great changes in pre-screening odds, DLR+ ≤10 and >5, and DLR− >0.1 and ≤0.2 indicate moderate changes, DLR+ ≤5 and >2, and DLR− >0.2 and ≤0.5 indicate small changes. DLR+ <2 and DLR− >0.5 are rarely important (Jaeschke et al., 1994).

Results

The results for research question 1 are shown in Table 1. The columns labeled bivariate analyses show that all subtests correlate significantly ($p < 0.001$) with spelling deficits, but the correlations are relatively small. The strongest correlation was found for sentence repetition ($r_{pb} = -0.236$) followed by letter knowledge ($r_{pb} = -0.218$) and initial phoneme detection ($r_{pb} = -0.203$). Children scoring higher on the sentence repetition test, the letter knowledge test, and the initial phoneme detection test at school entry were therefore less likely to show spelling deficits at the end of grade 1. The smallest correlation was found for the letter-number span forward ($r_{pb} = -0.116$). In line with the relatively low correlations, the AUC values for the single predictors are poor (i.e., <0.70). Thus, individually the subtests show relatively low diagnostic accuracy. To improve the diagnostic accuracy by using a set of important predictors simultaneously, we now focus on the results of the LASSO models, which are shown in the columns labeled LASSO models in Table 1. For the penalty term λ at which the AUC reaches its maximum, six out of the eight predictors emerge as important, with letter knowledge ($b = -0.389$) and sentence repetition ($b = -0.388$) being the strongest predictors. Initial phoneme detection ($b = -0.101$), RAN objects ($b = 0.138$), RAN digits ($b = 0.007$) and letter-number span backwards ($b = -0.079$) were also included in the prediction model. The more parsimonious λ_{AUC_1SE} model selected two predictors: sentence repetition ($b = -0.130$) and letter knowledge ($b = -0.076$). Further, we report the results of a standard logistic regression model with all predictors considered simultaneously for comparison (see Table 1, column logistic regression): These show that only letter knowledge ($b = -0.497$, $p < 0.01$) and sentence repetition ($b = -0.512$, $p < 0.01$) have significant effects on spelling deficits. However, when interpreting statistical significance, note that the collinearity of the predictors inflates the standard errors and therefore increases the p-values.

In regard to diagnostic accuracy (research question 2), we found that the six-predictor model resulted in an AUC of 0.736 [DeLong 95% CI (0.675; 0.797)], which suggests adequate diagnostic accuracy. The AUC for the two-predictor model was also fair and only slightly smaller than the AUC with six predictors [AUC = 0.725; DeLong 95% CI (0.666; 0.784)]. According to a bootstrapped test for paired ROC curves, the difference between the AUCs for λ_{AUC_max} and λ_{AUC_1SE} is not significant ($D = 0.919$, $p = 0.358$), which supports the use of the more parsimonious model with two predictors.

In order to provide information on the generalizability of the screening tool (research question 3), we report the results of analyses comparing ROC curves between groups. Table 2 provides AUC values for subgroups (girls vs. boys, and monolingual German-speaking children vs. children with another first language). The AUC based on the parsimonious λ_{AUC_1SE} model with three predictors was 0.746 [DeLong 95% CI (0.671; 0.821)] for boys and 0.692 [DeLong 95% CI (0.594; 0.791)] for girls. Both the bootstrapped test for unpaired ROC curves ($D = 0.855$, $p = 0.392$) and the Venkatraman test, which compares actual ROC curves ($E = 0.016$, $p = 0.475$), indicate that the ROC (AUC difference = 0.054) curves do not differ between boys and girls. Although the AUC difference between monolingual German-speaking children [AUC = 0.756; DeLong 95% CI (0.684; 0.829)] and children with a different first language [AUC = 0.658; DeLong 95% CI (0.540; 0.776)] is somewhat greater (AUC difference = 0.099), both tests again indicate that the ROC curves do not differ significantly ($D = 1.429$, $p = 0.153$; $E = 0.020$, $p = 0.207$). The same results (i.e., no significant differences in diagnostic accuracy between subgroups) were found for the λ_{AUC_max} screening score.

Finally, we evaluated various cut-offs for the screening scores (see Table 3). We report diagnostic accuracy statistics for cut-off values with sensitivity = 0.75, 0.80, 0.85, and 0.90. This range aims to capture various conditions (resources available to support students, etc.) at schools. A sensitivity of 0.75 results in a comparatively smaller number of screen-positives to be supported, at the cost of a high number of false negatives. For example, the parsimonious λ_{AUC_1SE} AUC screening score based on three predictors resulted, for a sensitivity of 0.75, in an empirically estimated SE of 0.750 [95% CI (0.644, 0.838)] and a rate of positive screening results of 42.3%. With this cut off, only 75% of the children with spelling deficits would be identified correctly as children in need of support. This cutoff yields a specificity of 0.639 [95% CI (0.592, 0.684)] and a PPV of 0.285 [95% CI (0.206, 0.408)], which implies that 28.5% of the positive screens finally show spelling deficits. The DLR+ of 2.079 [95% CI (1.745, 2.478)] and DLR− [0.391; 95% CI (0.268,

TABLE 1 Prediction of spelling deficits: bivariate and multivariate results.

	Bivariate analyses				Logistic regression			LASSO model	
		ROC-analysis						λ_{AUC_max}	λ_{AUC_1SE}
	r_{pb}	AUC	DeLong 95% CI		b	SE	p-value	b	b
Letter knowledge	−0.218***	0.677	0.616	0.738	**−0.497**	**0.164**	**0.002**	−0.389	−0.076
Initial phoneme detection	−0.203***	0.633	0.566	0.699	−0.145	0.145	0.319	−0.101	
Letter-number-span forward	−0.116**	0.606	0.537	0.676	0.149	0.146	0.307		
Letter-number-span backward	−0.193***	0.647	0.584	0.710	−0.124	0.146	0.395	−0.079	
RAN objects	0.175***	0.613	0.544	0.681	0.163	0.148	0.270	0.138	
RAN digits	0.169***	0.602	0.538	0.667	0.018	0.158	0.907	0.007	
Sentence repetition	−0.236***	0.683	0.622	0.744	**−0.512**	**0.182**	**0.005**	−0.388	−0.130
Receptive vocabulary	−0.178***	0.647	0.585	0.709	0.031	0.167	0.853		
Intercept					−1.915	0.146		−1.835	−1.661

TABLE 2 Comparison of ROC curves between subsamples.

	AUC	DeLong 95% CI		AUC difference	
λ_{AUC_max}				Value	Bootstrap (D)/Venkatraman (E)
Gender					
Boys	0.779	0.702	0.856	0.102	$D = 1.637$, $p = 0.102$
Girls	0.677	0.579	0.775		$E = 0.028$, $p = 0.103$
First language					
Monolingual German	0.760	0.688	0.832	0.079	$D = 1.110$, $p = 0.267$
Bilingual	0.681	0.557	0.804		$E = 0.016$, $p = 0.362$
Gender					
Boys	0.746	0.671	0.821	0.054	$D = 0.855$, $p = 0.392$
Girls	0.692	0.594	0.791		$E = 0.016$, $p = 0.475$
First language					
Monolingual German	0.756	0.684	0.829	0.099	$D = 1.429$, $p = 0.153$
Bilingual	0.658	0.540	0.776		$E = 0.020$, $p = 0.207$

0.570)] indicate small but important changes in the pre-screening odds. In contrast, a high sensitivity of 0.90 (i.e., a smaller number of false negatives) implies many screen-positives to be supported, which may overtax the resources available in schools. For the parsimonious λ_{AUC_1SE} AUC screening score, we found an empirically estimated SE of 0.905 [95% CI (0.821, 0.958)]. This cut off, however, resulted in 73.0% positive screening results, which means that about three out of four children would have to be supported in order to reach 90% of all children in need of support. Consequently, the specificity is very low at 0.306 [95% CI (0.263, 0.351)], and also the PPV [0.200; 95% CI (0.168, 0.375)] and the DLR+ [1.304; 95% CI (1.188, 1.431)] are low. Generally, the various cut-off values yield high proportions of students with a positive screening result (ranging from 43.5 to 72.3%) and thus low specificity values and PPVs, which indicate high rates of false positives.

Discussion

In international research, letter knowledge, oral skills (i.e., grammar and vocabulary), RAN, orthographic awareness, and phonological awareness have recently been identified as strong precursors of spelling (for a comprehensive review see Appendix 1). This is also the assumption of this study, however new is the for everyday school life suitable assessment procedure at the beginning to predict spelling at the end of the first year of primary school. By use of various precursor subtests, we designed a cost-free, tablet-based screening battery to predict written word spelling deficits in Austrian-German spelling. Particularly letter knowledge and grammar play a major part in this respect. The didactic implications of the screening variables letter knowledge and grammar are discussed below.

According to research question 1, we identified two models to predict German spelling ability in Austrian German: a broader model based on the variables letter knowledge, grammar, RAN digits, RAN objects, phoneme detection and phonological working memory. A second—even more robust—model highlights just the two most predictive variables, letter knowledge and grammar. Those predictors are in line with recent and prior research for German (Ennemoser et al., 2012; Fricke et al., 2016; Kastner-Koller et al., 2023). Other precursors, like RAN and phonological awareness are relevant but

TABLE 3 Diagnostic accuracy statistics for different cut-off values.

	Cut-off	Positive screens	SE	SP	PPV	NPV	DLR+	DLR−
λ_{AUC_max}								
SE = 0.75	0.141	50.9%	0.750 [0.644, 0.838]	0.539 [0.491, 0.586]	0.238 [0.205, 0.350]	0.918 [0.871, 0.932]	1.626 [1.386, 1.908]	0.464 [0.317, 0.679]
SE = 0.80	0.123	55.4%	0.798 [0.696, 0.877]	0.493 [0.445, 0.541]	0.232 [0.199, 0.354]	0.927 [0.881, 0.939]	1.574 [1.365, 1.814]	0.410 [0.266, 0.634]
SE = 0.85	0.106	62.8%	0.845 [0.750, 0.915]	0.411 [0.364, 0.459]	0.216 [0.184, 0.351]	0.933 [0.884, 0.944]	1.435 [1.272, 1.619]	0.377 [0.256, 0.629]
SE = 0.90	0.094	70.1%	0.905 [0.821, 0.958]	0.333 [0.293, 0.3807]	0.207 [0.175, 0.384]	0.948 [0.898, 0.957]	1.357 [1.233, 1.494]	0.286 [0.146, 0.560]
λ_{AUC_1SE}								
SE = 0.75	0.164	42.3%	0.750 [0.644, 0.838]	0.639 [0.592, 0.684]	0.285 [0.246, 0.408]	0.930 [0.889, 0.942]	2.079 [1.745, 2.478]	0.391 [0.268, 0.570]
SE = 0.80	0.160	47.7%	0.798 [0.696, 0.877]	0.571 [0.523, 0.618]	0.263 [0.229, 0.396]	0.936 [0.895, 0.947]	1.858 [1.595, 2.164]	0.355 [0.230, 0.546]
SE = 0.85	0.154	55.9%	0.857 [0.764, 0.924]	0.498 [0.450, 0.546]	0.247 [0.213, 0.399]	0.948 [0.907, 0.957]	1.706 [1.502, 1.939]	0.287 [0.169, 0.489]
SE = 0.90	0.145	73.0%	0.905 [0.821, 0.958]	0.306 [0.263, 0.351]	0.200 [0.168, 0.375]	0.944 [0.890, 0.954]	1.304 [1.188, 1.431]	0.311 [0.159, 0.611]

95% confidence intervals are in squared brackets.

lost importance in the given sample of first graders. Kastner-Koller et al. (2023) showed via the general school screening SES that a prediction of spelling at the beginning of second grade can be successful (42% predicted variance in writing one year after assessment in kindergarten). Here, working memory was identified as the most important predictor variable for reading, writing and arithmetic, followed by letter knowledge and phonological awareness. The authors proposed that phonological working memory is an underlying ability for different cognitive skills regarding speech.

For phonological awareness, we suppose that the specific predictive power of phonological awareness lies both in the younger sample age at the time of the initial survey (kindergarten) and in the small number of variables that were surveyed. Concerning the SES, grammar and vocabulary factors were not included. Furthermore, according to the authors (Kastner-Koller et al., 2023), the SES is decisive for predicting general school performance problems rather than specific performance deficits, such as difficulties in learning to write. Similar to the present study, Ennemoser et al. (2012) emphasized the importance of linguistic skills (i.e., grammar and vocabulary) for writing in the 2nd grade, whereas our results show the importance of grammar for subsequent writing as early as the 1st grade.

The predictive power of RAN for reading is a robust finding (see Landerl et al., 2022 for a review) and was recently shown in an Austrian sample of first graders (Schöfl et al., 2023a, 2023b). For writing, RAN also accounts for variance (Ennemoser et al., 2012; Kim et al., 2013; Lervåg and Hulme, 2010). Nevertheless, within our analyses model, RAN lost significance in predicting spelling at the end of first grade.

For research question 2, we found a reasonable diagnostic accuracy for a comprehensive and economical assessment battery. Other German batteries (Fricke et al., 2016) found comparable rates, for more details of effect sizes for spelling see Appendix 1.

As shown for research question 3, no significant differences in diagnostic accuracy were found between subgroups. There are no differences for boys and girls and for children with German as first or second language.

This result of missing gender disparities is in line with for example Dutch studies (Keuning and Verhoeven, 2008) for elementary school children from second grade on. For younger children, results are diverse, Allred (1990) found that girls from the first grade on are better spellers. For German language, recent studies on normative test batteries from first to seventh grade support our results of no gender disparities.

Concerning differences in spelling of mono- versus bilingual children, we found no significant group difference between mono- and bilinguals. We suppose that the missing differences are due to the provided vocabulary in the subtest. The words are based on typical word structures in German. Therefore, it is important to know these linguistically motivated structures on which the selected words are based. Knowledge about these structures could be independent of the children's individual language biography. To our knowledge further research is missing.

In summary the strongest predictors for word spelling skills at the end of first grade are letter knowledge and grammar. As there are no group differences concerning gender and language (children with German as L1 or L2), the didactic implications do consider the individual differences in spelling performance, but gender and/or language related disparities can be disregarded.

Didactic implications: letter knowledge

Early letter knowledge is a robust surface variable in spelling (Cordewener et al., 2012), comparable to rapid naming in reading

(Lervåg and Hulme, 2009). The strong relationship between letter knowledge and spelling skills might reflect an already existing access to written language itself. Children who know many letters might be able to show—at least implicitly—some elementary understanding of the basic phonological principle that is characteristic of alphabetical writing systems (Treiman, 2013).

Correlations between early letter knowledge and reading acquisition were described about 20 years ago (Foulin, 2005; Leppänen et al., 2008; Muter et al., 2004). Letter knowledge reflects seemingly implicit knowledge about the structure of written language. Simple training that focuses only on individual letters, however, is not effective for reading acquisition (in German: Scheerer-Neumann, 2006). Intervention studies instead support the linking hypothesis that it is the combination of letter knowledge, phonological exercises and subsequently the writing of syllables and words (Moraske et al., 2018). Isolated and exhaustive practice of letter knowledge at the preschool level is not a useful approach. Although the predictive power of phonological awareness is weaker than expected, the importance of PA for school exercises remains significant. In particular, the training of letters in combination with phonological awareness exercises should be emphasized (Moraske et al., 2018). To support children in this skill, Schnitzler (2008) suggested a variety of exercises, for instance, identifying words with the same initial phoneme or changing the initial phonemes in words. According to Mayer (2016), phonological awareness of smaller units, such as phonemes, develops during written language acquisition. In pratice, we suggest working with the initial phoneme at the beginning of written language acquisition rather than with various phoneme positions within the word which are more difficult to detect (Schnitzler, 2008). The combination of phonological awareness (i.e., detecting phonemes) and the basic alphabetic principle (i.e., letter-sound correspondence) is crucial for written language acquisition (Moraske et al., 2018). As the children are gradually introduced to letters and their names and sounds, initial phoneme detection is therefore to be systematically combined with the initial letter. This could be supported visually by the house-and-garage model mentioned above (Röber, 2013).

Regarding written language acquisition in grade 1, we infer that children with deficits in letter knowledge might benefit from integrated letter training that allows systematic insights not only into formal aspects, but also into aspects of the positioning of a letter within the syllable and the written canonical trochaic form of the German writing system. The letter <e> and its pronunciation depending on whether it is positioned in the strong or weak syllable might be a good example (Bredel, 2009; Fuhrhop and Peters, 2013). In practice, this could involve using use the house-and-garage model [originally designed for German trochaic words by Röber (2013)] to visualize for children the position of the latter within the syllable.

Didactic implications: grammar

For German spelling, oral grammar competence is an important precursor (von Goldammer et al., 2011), which is also confirmed by in the Schnapp spelling test. This might be explained by the strong grammatical (syllabic, morphological and syntactical) principles in the German writing system that act in concert. Theoretical approaches to grammar teaching and recent writing system research in German assume a strong relationship between grammar and literacy skills in general (Bredel, 2015b for an overview of the relationship between grammar and written language). Furthermore, the ability to recognize certain patterns in written language units is important and develops during written language acquisition. Dealing with written language has a number of prerequisite metalinguistic skills. Learning an alphabetical writing system then requires the abilities to identify sounds and to recognize how these sounds are transferred into written language. The ability to recognize certain patterns in written language units is also important (Haueis, 2015). Sufficient knowledge of grammar helps to detect syllable and morphological patterns in German word spelling and later, when dealing with sentences, to detect syntactic patters, such as space between words, German capitalization, and punctuation, when dealing with sentences (Bredel, 2015c). Morphological awareness is often cited as important for German word spelling (Görgen et al., 2021). Morphemes are units that are not only relevant for stem consistency, but also for identifying inflections. The latter is often placed in the weak *schwa* syllable. Stable access to stem consistency and how inflection is managed might improve spelling skills in general.

Grammar as a key precursor for German spelling adds weight to the argument that grammatical structures are very important in the German writing system. For children who show deficits when screened, knowledge about the canonical trochaic foot and related patterns allows them to gain insights into morphological principles far more easily (Röber, 2013; Bredel et al., 2017). Furthermore, working with syntactical structures, such as German V2-sentences, helps children to understand how sentences are constructed. To support language development in a practical way, children can be introduced to the topological field model (Granzow-Emden, 2015). For first graders, pictures or word cards could be used to form a spoken sentence instead of writing down sentences in the model and reading them aloud. Visualizing sentence structure promotes metalinguistic skills in children (Topalovic and Michalak, 2015).

Synopsis and limitations

The screening showed acceptable diagnostic accuracy [AUC = 0.757, DeLong 95% CI (0.698; 0.816)]. However, to achieve the desired sensitivity (>0.80), we must accept a relatively high rate of false positives (specificity = 0.57), which results in about one out of two pupils being screened positive. With a focus on intervention, this large group of screen-positives might thus challenge or even overtax teachers and school resources in providing support for these children. The problem of a high false-positive rate is not unique to our screening test (Cogo-Moreira et al., 2023; Duff et al., 2015; Johnson et al., 2009). Future research should therefore aim to improve specificity, for instance, by adding further subtests to the screening tool (Johnson et al., 2009). For example, reading socialization via parents or friends, motivational aspects, and further predictive child-related variables could be considered (only language-related variables were included in our approach). Specific attention processes or visual approaches were not considered, as the requisite data were not collected in this study. Regarding phonological awareness, only one subtest was included in the Schnapp screening battery. A wider approach to phonological awareness including both easier subtests, such as rhyming and syllable counting, and more difficult ones (e.g., manipulating phonemes) following the hierarchical structure of phonological tasks by Schnitzler (2008) should be analyzed.

Some researchers (Kim et al., 2013; Stainthorp et al., 2013) found orthographic awareness to be a significant precursor of early spelling skills. Actually, we do not include any task to tap orthographic awareness. Subsequent studies should add this variable for screening purposes.

Finally, continuous surveillance of spelling development is required because there may be various pathways to spelling difficulties. Many children with early difficulties do not go on to develop spelling problems, in the same way that many children who initially screen negative may later demonstrate spelling difficulties later, analogously to early screening for reading problems (Duff et al., 2015).

In summary our didactical implication to support written language acquisition for children in primary school is related to didactical implementations (Röber, 2013; Röber et al., 2019; Bredel, 2016; Bredel et al., 2017 based on the graphematical theory Primus, 2010; Evertz, 2016; Maas, 2022; Eisenberg, 2020; Fuhrhop and Peters, 2023)—both mentioned above. Children gain a comprehensive insight into the patterns of written language such as different sounds depending on letter positioning within the graphematic foot, how strong and weak syllables are presented in written language and how morphological stems are derivated from the graphematic foot, which is the initial point for constructing sentences.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Regional School Board for Upper Austria. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

MS: Conceptualization, Data curation, Investigation, Project administration, Writing – original draft, Writing – review & editing. GS: Conceptualization, Investigation, Project administration, Writing – original draft, Writing – review & editing. SZ: Writing – original draft, Writing – review & editing. CW: Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. All costs related to this research project were borne by the University of Education Upper Austria and Research Institute for Developmental Medicine (RID), Johannes Kepler University Linz. The tablets were provided by the project team with funds from the University of Education Upper Austria. Article processing charge was funded by the Johannes Kepler University Linz Open Access Publishing Fund.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2024.1378493/full#supplementary-material

References

Allred, R. A. (1990). Gender differences in spelling achievement in grades 1 through 6. *J. Educ. Res.* 83, 187–193. doi: 10.1080/00220671.1990.10885955

Bredel, U. (2009). Orthographie als System—Orthographieerwerb als Systemerwerb. *Z. Lit. Linguist.* 39, 135–155. doi: 10.1007/BF03379820

BMBWF. (2024). Förderorientiertes Schuleingangsscreenin. Available at: https://ses.bmbwf.gv.at (Accessed October 9, 2024).

Bredel, U. (2015a). "Der Aufbau von elementarem Wissen über die Systematik der Orthographie deutscher Wörter" in Schriftsprach- und Orthographieerwerb: Erstlesen, Erstschreiben. eds. C. Röber and H. Olfert (Baltmannsweiler: Schneider Verlag Hohengehren), 255–279.

Bredel, U. (2015b). "Grammatik und Literalität" in Deutschunterricht in Theorie und Praxis: Band 6. Deutschunterricht in Theorie und Praxis: Sprachreflexion und Grammatikunterricht. ed. H. Gornik. *2. korrigierte Auflage* ed (Baltmannsweiler: Schneider Verlag Hohengehren GmbH), 160–173.

Bredel, U. (2015c). "Sprachreflexion und Orthographie" in Deutschunterricht in Theorie und Praxis: Band 6. Deutschunterricht in Theorie und Praxis: Sprachreflexion und Grammatikunterricht. ed. H. Gornik. *2. korrigierte Auflage* ed (Baltmannsweiler: Schneider Verlag Hohengehren GmbH), 266–281.

Bredel, U. (2016). "23. Schriftspracherwerb" in Handbuch Laut, Gebärde, Buchstabe. eds. U. Domahs and B. Primus (Berlin: De Gruyter), 436–454.

Bredel, U., Fuhrhop, N., and Noack, C. (2017). Wie Kinder lesen und schreiben lernen. Tübingen: Narr Francke Attempo Verlag GmbH + Co. KG.

Breit, S., Bruneforth, M., and Schreiner, C. (2016). Standardüberprüfung Deutsch, 4. Schulstufe. Salzburg: BMBWF.

Caravolas, M., Hulme, C., and Snowling, M. J. (2001). The foundations of spelling ability: evidence from a 3-year longitudinal study. *J. Mem. Lang.* 45, 751–774. doi: 10.1006/jmla.2000.2785

Cogo-Moreira, H., Molinari, G. L., Carvalho, C. A. F., Kida, A. S. B., Lúcio, P. S., and Avila, C. R. B. (2023). Cut-off point, sensitivity and specificity for screening the reading fluency in children. *CoDAS* 35:e20210263. doi: 10.1590/2317-1782/20232021263en

Cordewener, K. A. H., Bosman, A. M. T., and Verhoeven, L. (2012). Predicting early spelling difficulties in children with specific language impairment: a clinical perspective. *Res. Dev. Disabil.* 33, 2279–2291. doi: 10.1016/j.ridd.2012.07.003

Denckla, M. B., and Rudel, R. (1974). Rapid "automatized" naming of pictured objects, colors, letters and numbers by normal children. *Cortex* 10, 186–202. doi: 10.1016/s0010-9452(74)80009-2

Duff, F. J., Mengoni, S. E., Bailey, A. M., and Snowling, M. J. (2015). Validity and sensitivity of the phonics screening check: implications for practice. *J. Res. Read.* 38, 109–123. doi: 10.1111/1467-9817.12029

Eisenberg, P. (1989). "Die Schreibsilbe im Deutschen" in Schriftsystem und Orthographie. eds. P. Eisenberg and H. Günther (Berlin: De Gruyter), 57–84.

Eisenberg, P. (2020). Grundriss der deutschen Grammatik. Berlin: J.B. Metzler Stuttgart.

Elliott, J. (2018). "Family engagement in urban education: the impact and factors affecting engagement" in Master of Education Applied Research Projects, 5. Available at: https://digitalcommons.cedarville.edu/education_research_projects/5

Enders, C. K. (2010). Applied missing data analysis. New York: Guilford Press.

Endlich, D., Küspert, P., Lenhard, W., Marx, P., and Schneider, W. (2019). Laute, Reime, Sprache—Würzburger Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten. (Berlin: De Gruyter).

Ennemoser, M., Marx, P., Weber, J., and Schneider, W. (2012). Spezifische Vorläuferfertigkeiten der Lesegeschwindigkeit, des Leseverständnisses und des Rechtschreibens. *Z. Entwicklungspsychologie Padagog. Psychol.* 44, 53–67. doi: 10.1026/0049-8637/a000057

Evertz, M. (2016). "20. Graphematischer Fuß und graphematisches Wort" in Handbuch Laut, Gebärde, Buchstabe. eds. U. Domahs and B. Primus (De Gruyter), 377–397.

Foulin, J. N. (2005). Why is letter-name knowledge such a good predictor of learning to read? *Read. Writ.* 18, 129–155. doi: 10.1007/s11145-004-5892-2

Fricke, S., Szczerbinski, M., Fox-Boyer, A., and Stackhouse, J. (2016). Preschool predictors of early literacy acquisition in German-speaking children. *Read. Res. Q.* 51, 29–53. doi: 10.1002/rrq.116

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33, 1–22. doi: 10.18637/jss.v033.i01

Friedman, J., Hastie, T., Tibshirani, R., Narasimhan, B., Tay, K., Simon, N., et al. (2021). glmnet: LASSO and elastic-net regularized generalized linear models. Heidelberg: J.B. Metzlersche Verlagsbuchhandlung. R Core Team.

Frith, U. (1986). A developmental framework for developmental dyslexia. *Ann. Dyslexia* 36, 67–81. doi: 10.1007/BF02648022

Fuhrhop, N., and Peters, J. (2013). Einführung in die Phonologie und Graphematik: J.B. Metzler Stuttgart.

Fuhrhop, N., and Peters, J. (2023). Einführung in die Phonologie und Graphematik. *2nd* Edn: Baltmannsweiler: Schneider Verlag Hohengehren. J.B. Metzler Stuttgart.

González-Valenzuela, M.-J., Lopez-Montiel, D., Díaz-Giraldez, F., and Martín-Ruiz, I. (2023). Word writing and cognitive predictors in Spanish at the age of seven. *Int. J. Early Child.* doi: 10.1007/s13158-023-00359-w

Görgen, R., De Simone, E., Schulte-Körne, G., and Moll, K. (2021). Predictors of reading and spelling skills in German: the role of morphological awareness. *J. Res. Read.* 44, 210–227. doi: 10.1111/1467-9817.12343

Granzow-Emden, M. (2015). "Sprachstrukturen verstehen: Die Entwicklung grammatischer Kategorien" in Deutschunterricht in Theorie und Praxis: Band 6. Deutschunterricht in Theorie und Praxis: Sprachreflexion und Grammatikunterricht. ed. H. Gornik. *2. korrigierte Auflage* ed (Schneider Verlag Hohengehren GmbH), 213–241.

Grob, A., Meyer, C., and Hagmann-von Arx, P. (2009). Intelligence and Development Scales (IDS). Intelligenz- und Entwicklungsskalen für Kinder von 5–10 Jahren. Bern: Hans Huber.

Haueis, E. (2015). "Sprachreflexion und Schriftspracherwerb" in Deutschunterricht in Theorie und Praxis: Band 6. Deutschunterricht in Theorie und Praxis: Sprachreflexion und Grammatikunterricht. ed. H. Gornik. *2. korrigierte Auflage* ed (Baltmannsweiler: Schneider Verlag Hohengehren GmbH), 184–193.

Henschel, S., Heppt, B., and Weirich, S. (2023). "Zuwanderungsbezogene Disparitäten" in IQB-Bildungstrend 2022: Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im dritten Ländervergleich. eds. P. Stanat, S. Schipolowski, R. Schneider, S. Weirich, S. Henschel and K. A. Sachse (Münster: Waxmann), 299–344. Waxmann

Ibrahim, L. A., Hamann, C., and Öwerdieck, D. (2018). Identifying specific language impairment (SLI) across different bilingual populations: German sentence repetition task (SRT). In A. B. Bertolini and M. J. Kaplan (Eds.), Proceedings of the 42nd annual Boston University Conference on Language Development. 1–14. Cascadilla Press.

Jaeschke, R., Guyatt, G. H., and Sackett, D. L. (1994). Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The evidence-based medicine working group. *JAMA* 271, 703–707. doi: 10.1001/jama.1994.03510330081039

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). An introduction to statistical learning, vol. *103*. New York: Springer.

Jansen, H., Mannhaupt, G., Marx, H., and Skowronek, H. (1999). Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (BISC). Göttingen: Hogrefe.

Jöbstl, V., Steiner, A. F., Kastner-Koller, U., Deimann, P., Kaltenberger, A., Wagner, V., et al. (2022). Entwicklung eines förderorientierten Schuleingangsscreenings: Erste Befunde zur prognostischen Validität. *Frühe Bildung* 11, 194–200. doi: 10.1026/2191-9186/a000585

Johnson, E. S., Jenkins, J. R., Petscher, Y., and Catts, H. W. (2009). How can we improve the accuracy of screening instruments? *Learn. Disabil. Res. Pract.* 24, 174–185. doi: 10.1111/j.1540-5826.2009.00291.x

Kastner-Koller, U., Kaltenberger, A., Jöbstl, V., Landerl, K., Steiner, A. F., Wagner, V., et al. (2023). Zur Vorhersage des schulischen Lern- und Arbeitsverhaltens durch domänenübergreifende Vorläuferfähigkeiten. *Lernen Und Lernstörungen* 12, 115–125. doi: 10.1024/2235-0977/a000412

Kessler, B., Pollo, T., Treiman, R., and Cardoso-Martins, C. (2013). Frequency analyses of prephonological spellings as predictors of success in conventional spelling. *J. Learn. Disabil.* 46, 252–259. doi: 10.1177/0022219412449440

Keuning, J., and Verhoeven, L. (2008). Spelling development throughout the elementary grades: the Dutch case. *Learn. Individ. Differ.* 18, 459–470. doi: 10.1016/j.lindif.2007.12.001

Kim, Y.-S., Apel, K., and Al Otaiba, S. (2013). The relation of linguistic awareness and vocabulary to word reading and spelling for first-grade students participating in response to intervention. *Lang. Speech Hear. Serv. Sch.* 44, 337–347. doi: 10.1044/0161-1461(2013/12-0013)

Landerl, K., Castles, A., and Parrila, R. (2022). Cognitive precursors of reading: a cross-linguistic perspective. *Sci. Stud. Read.* 26, 111–124. doi: 10.1080/10888438.2021.1983820

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., et al. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *J. Child Psychol. Psychiatry* 54, 686–694. doi: 10.1111/jcpp.12029

Lenhart, J., Marx, P., Segerer, R., and Schneider, W. (2019). Rechtschreibung ohne Schreiben. *Diagnostica* 65, 216–227. doi: 10.1026/0012-1924/a000229

Leppänen, U., Aunola, K., Niemi, P., and Nurmi, J.-E. (2008). Letter knowledge predicts grade 4 reading fluency and reading comprehension. *Learn. Instr.* 18, 548–564. doi: 10.1016/j.learninstruc.2007.11.004

Lervåg, A., and Hulme, C. (2009). Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychol. Sci.* 20, 1040–1048. doi: 10.1111/j.1467-9280.2009.02405.x

Lervåg, A., and Hulme, C. (2010). Predicting the growth of early spelling skills: are there heterogeneous developmental trajectories? *Sci. Stud. Read.* 14, 485–513. doi: 10.1080/10888431003623488

López-Ratón, M., Rodríguez-Álvarez, M. X., Suárez, C. C., and Sampedro, F. G. (2014). OptimalCutpoints: AnRPackage for selecting optimal cutpoints in diagnostic tests. *J. Stat. Softw.* 61, 1–36. doi: 10.18637/jss.v061.i08

Maas, U. (1992). Grundzüge der deutschen Orthographie. *De Gruyter.* Berlin: De Gruyter. doi: 10.1515/9783111376974

Maas, U. (2022). "Laute und Buchstaben—zu den phonographischen Grundlagen des Schrifterwerbs" in Deutschunterricht in Theorie und Praxis: Vol. 2. Schriftsprach- und Orthographieerwerb: Erstlesen, Erstschreiben. eds. C. Röber and H. Olfert, Baltmannsweiler: Schneider Verlag Hohengehren GmbH.113–139.

Martschinke, S., Kirschhock, E., and Frank, A. (2001). Der Rundgang durch Hörhausen: Erhebungsverfahren zur phonologischen Bewusstheit (1. Aufl.). Diagnose und Förderung im Schriftspracherwerb: Bd. 1. Donauwörth: Auer.

Marx, P., and Lenhard, W. (2010). "Diagnostische Merkmale von Screeningverfahren" in Frühprognose schulischer Kompetenzen. eds. M. Hasselhorn and W. Schneider (Göttingen: Hogrefe), 68–84.

Mayer, A. (2016). Lese-Rechtschreibstörungen (LRS). München, Basel: Ernst Reinhardt Verlag.

McNeish, D. M. (2015). Using LASSO for predictor selection and to assuage overfitting: a method long overlooked in behavioral sciences. *Multivar. Behav. Res.* 50, 471–484. doi: 10.1080/00273171.2015.1036965

Moraske, S., Wyschkon, A., Poltz, N., Kucian, K., von Aster, M., and Esser, G. (2018). LRS-Prävention bei Risikokindern: Langfristige Effekte bis in die 3. Klasse. *Lernen Und Lernstörungen* 7, 171–183. doi: 10.1024/2235-0977/a000213

Muter, V., Hulme, C., Snowling, M. J., and Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development:

evidence from a longitudinal study. *Dev. Psychol.* 40, 665–681. doi: 10.1037/0012-1649.40.5.665

Nag, S., Vagh, S. B., Dulay, K. M., and Snowling, M. J. (2019). Home language, school language and children's literacy attainments: a systematic review of evidence from low- and middle-income countries. *Rev. Educ.* 7, 91–150. doi: 10.1002/rev3.3130

Niemietz, J., Jindra, C., Schneider, R., Schumann, K., Schipolowski, S., and Sachse, K. A. (2023). "Soziale Disparitäten" in IQB-Bildungstrend 2022: Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im dritten Ländervergleich. eds. P. Stanat, S. Schipolowski, R. Schneider, S. Weirich, S. Henschel and K. A. Sachse (Paris: OECD Publishing), 261–298. doi: 10.1787/53f23881-en

Niolaki, G. Z., Papadimitriou, V., Terzopoulos, A. R., and Masterson, J. (2024). Greek-spelling predictors; an investigation of literacy- and cognitive-related factors. *J. Cogn. Psychol.* 36, 378–393. doi: 10.1080/20445911.2024.2322766

Niolaki, G. Z., Vousden, J., Terzopoulos, A. R., Taylor, L. M., Sephton, S., and Masterson, J. (2020). Predictors of single word spelling in English speaking children: a cross sectional study. *J. Res. Read.* 43, 577–596. doi: 10.1111/1467-9817.12330

OECD (2023). PISA 2022 results (volume I): the state of learning and equity in education: OECD Publishing.

Pan, S. C., Rickard, T. C., and Bjork, R. A. (2021). Does spelling still matter—and if so, how should it be taught? *Educ Psychol Rev* 33, 1523–1552. doi: 10.1007/s10648-021-09611-y

Pepe, M. S. (2004). The statistical evaluation of medical tests for classification and prediction: Oxford University Press.

Primus, B. (2010). "Strukturelle Grundlagen des deutschen Schriftsystems" in Schriftsystem und Schrifterwerb: linguistisch—didaktisch—empirisch. eds. U. Bredel, A. Müller and G. Hinney (Berlin: Walter de Gruyter GmbH & Co. KG), 9–45.

Röber, C. (2013). Die Leistungen der Kinder beim Lesen- und Schreibenlernen. Grundlagen der Silbenanalytischen Methode: Ein Arbeitsbuch mit Übungsaufgaben. Hohengehren: Schneider-Verlag Hohengehren.

Röber, C., Häusle, R., and Berchtold, M. (2019). "Erstklässler entdecken die Orthographie als Hilfe für das Rechtlesen und Rechtschreiben: Vorstellung des Schriftsprachstrukturierenden Konzeptes PALOPE" in Forschung & Entwicklung. Feldkirch: Forschungszeitschrift. 47–60.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:77. doi: 10.1186/1471-2105-12-77

Schaumburg, H. (2015). Chancen und Risiken digitaler Medien in der Schule: Medienpädagogische und -didaktische Perspektiven. Berlin: Bertelsmann Stiftung.

Scheerer-Neumann, G. (2006). "Entwicklung der basalen Lesefähigkeit" in Didaktik der deutschen Sprache—Band 1. eds. U. Bredel, H. Günther, P. Klotz, J. Ossner and G. Siebert-Ott (Stuttgart, Deutschland: UTB GmbH), 513–524.

Schnitzler, C. (2008). Phonologische Bewusstheit und Schriftspracherwerb: 31 Tabellen. Forum Logopädie. Stuttgart, New York, NY: Thieme.

Schöfl, M., Seifert, S., Steinmair, G., and Weber, C. (2023a). Are both object and alphanumeric rapid automatized naming measures required to predict word Reading fluency in German prereaders? *J. Psychoeduc. Assess.* 41, 811–827. doi: 10.1177/07342829231185375

Schöfl, M., Steinmair, G., Holzinger, D., and Weber, C. (2022). Predicting word reading deficits using an app-based screening tool at school entry. *Front. Pediatr.* 10:863477. doi: 10.3389/fped.2022.863477

Schöfl, M., Steinmair, G., Zepnik, S., and Gnambs, T. (2023b). Entwicklung eines digitalen Rechtschreibtests für die erste Klasse Grundschule: Dimensionalität und Reliabilität des SCHNAPP-Rechtschreibtests. *Lernen Und Lernstörungen* 12, 188–200. doi: 10.1024/2235-0977/a000404

Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., and Kliegl, R. (2015). childLex–Eine lexikalische Datenbank zur Schriftsprache für Kinder im Deutschen. *Psychol. Rundsch.* 66, 155–165. doi: 10.1026/0033-3042/a000275

Seifert, S., Paleczek, L., and Gasteiger-Klicpera, B. (2019). Rezeptive Wortschatzleistungen in der Grundschule: Unterschiede zwischen Kindern mit Deutsch als Erst- und Zweitsprache. *Empirische Sonderpädagogik* 4, 259–278. doi: 10.25656/01:18334

Silverman, R. D., McNeish, D., Speece, D. L., and Ritchey, K. D. (2021). Early screening for decoding- and language-related reading difficulties in first and third grades. *Assess. Eff. Interv.* 46, 99–109. doi: 10.1177/1534508419857234

Sirin, S. R. (2005). Socioeconomic status and academic achievement: a meta-analytic review of research. *Rev. Educ. Res.* 75, 417–453. doi: 10.3102/00346543075003417

Stainthorp, R., Powell, D., and Stuart, M. (2013). The relationship between rapid naming and word spelling in English. *J. Res. Read.* 36, 371–388. doi: 10.1111/jrir.12002

Statistik Austria. (2024). STATcube—Statistische Datenbank von STATISTIK AUSTRIA. Available at: https://statcube.at/

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science* 240, 1285–1293. doi: 10.1126/science.3287615

Thomé, G. (2006). "Entwicklung der basalen Rechtschreibkenntnisse" in Didaktik der deutschen Sprache—Band 1. eds. U. Bredel, H. Günther, P. Klotz, J. Ossner and G. Siebert-Ott (Stuttgart, Deutschland: UTB GmbH), 369–379.

Thompson, R. A. (2015). "Relationships, regulation, and early development" in Handbook of child psychology and developmental science: socioemotional processes. eds. M. E. Lamb and R. M. Lerner. *7th* ed (Oxford: John Wiley & Sons), 1–246. doi: 10.1002/9781118963418

Topalovic, E., and Michalak, M. (2015). "Sprachreflexion und Grammatik zwischen DaM und DaZ" in Grundlagen der Sprachdidaktik Deutsch als Zweitsprache. eds. M. Michalak and M. Kuchenreuther. *3. überarbeitete Auflage* ed (Baltmannsweiler: Schneider Verlag Hohengehren GmbH), 226–250.

Treiman, R. (2013). "Knowledge about letters as a foundation for reading and spelling" in Handbook of orthography and literacy. eds. R. M. Joshi and P. G. Aaron (New York: Routledge).

Treiman, R., and Bourassa, D. (2000). Children's written and oral spelling. *Appl. Psycholinguist.* 21, 183–204. doi: 10.1017/S0142716400002022

Treiman, R., Hulslander, J., Olson, R. K., Samuelsson, S., Elwér, Å., Furnes, B., et al. (2023). Predicting later spelling from kindergarten spelling in U.S., Australian, and Swedish children. *Reading* 27, 428–442. doi: 10.1080/10888438.2023.2186234

Treiman, R., Hulslander, J., Olson, R. K., Willcutt, E. G., Byrne, B., and Kessler, B. (2019). The unique role of early spelling in the prediction of later literacy performance. *Sci. Stud. Read.* 23, 437–444. doi: 10.1080/10888438.2019.1573242

Treiman, R., and Kessler, B. (2004). The case of case: children's knowledge and use of upper- and lowercase letters. *Appl. Psycholinguist.* 25, 413–428. doi: 10.1017/S0142716404001195

Treiman, R., Kessler, B., Decker, K., and Pollo, T. C. (2016). How do prephonological writers link written words to their objects? *Cogn. Dev.* 38, 89–98. doi: 10.1016/j.cogdev.2016.02.002

Venkatraman, E. S. (2000). A permutation test to compare receiver operating characteristic curves. *Biometrics* 56, 1134–1138. doi: 10.1111/j.0006-341x.2000.01134.x

von Goldammer, A., Mähler, C., Bockmann, A. K., Hasselhorn, M., and Hasselhorn, M. (2010). Vorhersage früher Schriftsprachleistungen aus vorschulischen Kompetenzen der Sprache und der phonologischen Informationsverarbeitung. *Z. Entwicklungspsychologie Padagog. Psychol.* 42, 48–56. doi: 10.1026/0049-8637/a000005

von Goldammer, A., Mähler, C., Hasselhorn, M., and Hasselhorn, M. (2011). Determinanten von Satzgedächtnis-Leistungen bei deutsch- und mehrsprachigen Vorschulkindern. *Z. Entwicklungspsychologie Padagog. Psychol.* 43, 1–15. doi: 10.1026/0049-8637/a000028

Wagner, R. K., and Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol. Bull.* 101, 192–212. doi: 10.1037/0033-2909.101.2.192

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychol. Bull.* 91, 461–481. doi: 10.1037/0033-2909.91.3.461

Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: we are ready to ROC. *J. Pediatr. Psychol.* 39, 204–221. doi: 10.1093/jpepsy/jst062

Yulianti, K., Denessen, E., and Droop, M. (2023). The effects of parental involvement on children's education. *Int. J. Parents Educ.* 10, 14–32. doi: 10.54195/ijpe.14123

Zarić, J., Hasselhorn, M., and Nagler, T. (2020). Orthographic knowledge predicts reading and spelling skills over and above general intelligence and phonological awareness. *Psychol. Educ.* 36, 21–43. doi: 10.25656/01:22749