Journal of Research in Reading
UKLA The United Kingdom Literacy Association
Journal of Research in Reading, ISSN 0141-0423
Volume 44, Issue 4, 2021, pp 817–836
DOI:10.1111/1467-9817.12373

Influences from working memory, word and sentence reading on passage comprehension and teacher ratings

Nicole Kirschmann
Department of Psychology, University of Wuerzburg, Wuerzburg, Germany

Wolfgang Lenhard
Department of Psychology, University of Wuerzburg, Wuerzburg, Germany

Sebastian Suggate
University of Regensburg, Regensburg, Germany

Reading fluency is a major determinant of reading comprehension but depends on moderating factors such as auditory working memory (AWM), word recognition and sentence reading skills. We investigated how word and sentence reading skills relate to reading comprehension differentially across the first 6 years of schooling and tested which reading variable best predicted teacher judgements. We conducted our research in a rather transparent language, namely, German, drawing on two different data sets. The first was derived from the normative sample of a reading comprehension test (ELFE-II), including 2056 first to sixth graders with readings tests at the word, sentence and text level. The second sample included 114 students from second to fourth grade. The latter completed a series of tests that measured word and sentence reading fluency, pseudoword reading, AWM, reading comprehension, self-concept and teacher ratings. We analysed the data via hierarchical regression analyses to predict reading comprehension and teacher judgements. The impact of reading fluency was strongest in second and third grade, afterwards superseded by sentence comprehension. AWM significantly contributed to reading comprehension independently of reading fluency, whereas basic decoding skills disappeared after considering fluency. Students' AWM and reading comprehension predicted teacher judgements on reading fluency. Reading comprehension judgements depended both on the students' self-concept and reading comprehension. Our results underline that the role of word reading accuracy for reading comprehension quickly diminishes during elementary school and that teachers base their assessments mainly on the current reading comprehension skill.

Keywords: reading skills, reading comprehension, teacher assessments, passage comprehension, word recognition

Highlights

What is already known about this topic

- Reading fluency, auditory working memory and basic word recognition skills are predictors of reading comprehension.
- The orthographic depth of a language influences the acquisition of reading skills.
- Teacher judgements correspond moderately with actual students' performances.

What this paper adds

- In German, sentence reading fluency is a better predictor of reading comprehension than word reading fluency, from the first grades onwards.
- Basic word recognition skills did not predict reading comprehension in German from the second grade onwards.
- Self-concept, working memory and reading comprehension predict teacher judgements concerning reading fluency and comprehension.

Implications for theory, policy or practice

- The association between accurate word decoding and comprehension decreases across the first grade in transparent orthographies.
- Fluency at the word and sentence level best predicts reading comprehension and should be preferred over accuracy measures, at least in German.

In order to convert code into language and ideas, the ability to read fluently is essential for reading comprehension (Therrien, 2004). Beginning readers first acquire the ability to match phonemes onto graphemes at a syllable and word level, before starting to read longer sentences and thereafter passages of connected text (Adams, 1990). This process is dependent on a number of additional factors, specifically working memory, language transparency and children's reading skill. Reading tests typically focus on decoding, word fluency and passage comprehension; however, sentence reading skill may be a conceptually relevant predictor of reading comprehension, particularly in languages such as German that contain more regular sound-to-spelling mappings. Research comparing the relative contributions of these factors in German is scarce, with the same applying to associations with self-concept of reading and teacher judgements.

Reading fluency at word and sentence levels as a determinant of reading comprehension

Theories posit a clear association between word reading fluency and reading comprehension, whereby the former is necessary for the latter (e.g., Hoover & Gough, 1990; Kim, Wagner, & Lopez, 2012; Schwanenflugel et al., 2006). Reading fluency refers to skill at rapidly decoding lexical items in graphemic form, thereby generating semantic

information, either out loud or internally (Therrien, 2004). Reading comprehension refers, in turn, to extracting semantic information from text, otherwise known as reading for meaning (Keenan, Betjemann, & Olson, 2008). At a theoretical level, the Simple View of Reading (Gough & Tunmer, 1986) posits that reading comprehension emerges as the product of word recognition skill and language comprehension. The Simple View hence reflects the idea that word recognition skills alone do not suffice for reading for meaning, with other factors such as working memory playing a role (Joshi, Tao, Aaron, & Quiroz, 2012). At a word recognition level, the dual route cascade model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) points to the different skills required in word reading fluency, depending on whether this is automated or not. Thus, when reading unfamiliar or irregular words, readers have to effortfully decode, often letter by letter, until the word is sounded out, or a match with a lexical entry is obtained – which also places a demand on working memory (Leong, Tse, & Hau, 2008).

Studies typically investigate reading fluency using word reading (e.g., Torgesen, Wagner, & Rashotte, 2012) or passage reading tasks. To test fluency and mastery of the code, nonword or pseudoword reading tasks are often used because these partial out the influence of prior lexical knowledge in solving word items (Herman, Matyas, & Pratt, 2006). Word reading development represents a shift from more basic decoding to higher order, semantic related process, which takes place during the first years of reading and spelling acquisition. Specifically, this shift emerges from pre-alphabetic stages via the identification of contextual clues, via alphabetic processes matching graphemes to phonemes, moving to automated sight word reading with access to semantics (Ehri, 2005).

Because word recognition skills represent a bottleneck for reading comprehension in beginning readers, their effect should be most pronounced in the first grades. Meta-analyses report correlations of up to $r = .86$ for young readers and a decrease in this correlation to $r = .41$ in adults, with a pronounced turnaround approximately at the age of 10 (García & Cain, 2014). With respect to higher order processes that mirror automaticity, findings focusing on sentence reading fluency are less common and consistent in the English language. For example, Jenkins, Fuchs, van den Broek, Espin, and Deno (2003) found that sentence reading fluency in fourth grade was a significant predictor of reading comprehension after controlling for word reading fluency. However, Schwanenflugel et al. (2006) observed that reading fluency at sentence level no longer significantly related to reading comprehension in first, second and third graders, when word reading fluency was statistically controlled. These two observations suggest that the influence of sentence reading fluency could change across reading development. In fact, this assumption is supported by various studies (e.g., Kim et al., 2012; Klauda & Guthrie, 2008).

For writing systems such as German with more regular sound-to-spelling mappings, surprisingly little research has examined the comparative role that word and sentence reading fluency play in reading comprehension. For example, in German speaking countries, the research tradition has focused on assessing the influence of pre-reading skills and determinants such as phonemic awareness, IQ, working memory and oral language (Ennemoser, Marx, Weber, & Schneider, 2012; von Goldammer et al., 2010; Näslund, 1990). These findings highlight the importance of pre-schoolers' IQ, working memory, phonemic awareness and rapid naming for decoding speed at the end of first grade, which in turn, in combination with vocabulary and syntax, longitudinally predict reading comprehension years later. Kim et al.'s (2012) findings that word reading fluency is uniquely related to reading comprehension in the first, but not second grade, and that sentence reading fluency is uniquely related to reading comprehension in the second, but

not first grade, are thus particularly interesting because the subjects were children whose native language was Korean. Modern forms of the Korean writing systems feature 24 graphemes, and although not being completely transparent, the orthography is substantially more regular than English – actually approximating German in terms of orthographic depth. This sound-to-spelling regularity might moderate the effects of word versus sentence reading fluency on reading comprehension. Reading in irregular orthographies in contrast requires the exact identification of sublexical units like morphemes for reading comprehension (Seymour, Aro, & Erskine, 2003). Additionally, previous research has shown that vocabulary plays a greater role in decoding words in non-transparent languages (Suggate, Reese, Lenhard, & Schneider, 2014). At a theoretical level, individual word reading might rely more on higher order language processes, such as vocabulary and reading comprehension, due to the absence of contextual clues. Accordingly, sentence, and particularly passage reading, provides more contextual information to facilitate decoding of irregular words.

Factors moderating the relationship between fluency and comprehension

Besides decoding, there are a number of cognitive determinants of comprehension (Joshi et al., 2012), with auditory working memory (AWM) thought to play an important role in decoding (Cain, Oakhill, & Bryant, 2004). This seems plausible because graphemes have to be transcoded into their phonological form and stored in memory in order to be blended into words, at least on the indirect route to word recognition and most pronounced at the beginning of reading and spelling acquisition. Accordingly, a clear relationship between AWM and reading fluency has been established (Kibby, Lee, & Dyer, 2014; Siu, McBride, Tse, Tong, & Maurer, 2018). However, AWM is not only needed for transcoding graphemes and phonemes: In order to understand entire texts, there has to be a representation of the phonetic-lexical propositions of semantic content. Applying this to the Simple View of Reading (Gough & Tunmer, 1986), the AWM could therefore be not only a predictor of decoding, but also of reading comprehension (Seigneuric & Ehrlich, 2005). AWM thus plays a role in building integrated mental representations, which in turn, facilitate setting up situational models and drawing inferences (Cain et al., 2004).

In support of the role of AWM in reading comprehension, Wagner, Torgesen, and Rashotte (1994) found medium to high correlations between the phonological loop and reading comprehension. Studies show not only a direct path between working memory and reading comprehension but also an indirect one via phonological processing skills (e.g., Cain et al., 2004; Leather & Henry, 1994). Even after controlling for decoding and comprehension (e.g., vocabulary), a significant influence of AWM on reading comprehension remained. In sum, AWM seems to exert an influence both on reading fluency as well as on reading comprehension.

Subjective performance assessments by teachers: Precision and determinants

As shown, there are many standardised predictors influencing literacy development (Cromley & Azevedo, 2007). Given the importance of feedback in learning (e.g., Hattie, 2012) teacher judgement is likely central because of its immediacy and value to students (Südkamp, Kaiser, & Möller, 2012). Further, teacher judgements flow into remedial and instructional decisions (Zhu, Urhahne, & Rubie-Davies, 2018) but are often

not based on objective reading measures, which leads to questions about their incremental predictive validity.

On the one hand, there are findings suggesting that teachers' are only slightly consistent with the students' objective performances (e.g., Karing, Matthäi, & Artelt, 2011). On the other hand, Hoge and Coladarci (1989) report correlations of $r = .66$ between objectively measured student performance and subjective teacher judgement. Specific studies on language arts grades show a significant influence of students' reading performances on global and task-specific teacher judgements, finding correlations up to $r = .70$ (e.g., Feinberg & Shapiro, 2003; Rausch, 2016). In turn, teacher judgement explains additional variance in reading performance above and beyond word decoding and student motivation (van Kraayenoord & Schneider, 1999).

Although the actual reading skills of students play an important role, teacher judgements could however be influenced by more factors than the students' performance, such as sex, ethnic status and behaviours (Kaiser, Südkamp, & Möller, 2017), calling the validity of teacher judgements partly into question (Feinberg & Shapiro, 2009). As general academic predictors like AWM have an impact on the reading performance (Kibby et al., 2014; Siu et al., 2018), it seems clear that they might also influence teacher judgements. The same applies to the students' self-concepts, which reflect cognitions and appraisals of children's own reading achievement (Segerer, Niklas, Suggate, & Schneider, 2020). Consequently, a student's reading self-concept could also directly influence the teacher's judgement (Kriegbaum, Steinmayr, & Spinath, 2019), when the student's overt attitude towards reading influences the teacher's assessment of the actual performance. Conceivably, teacher judgement might be influenced by different dimensions of children's reading skills. It might be expected that teachers receive information on children's development through hearing them read aloud in class, which should be especially true for younger readers, and via their performance in subjects relying on reading comprehension skills for the acquisition of knowledge. Given that word reading is used more in assessment than in educational practice, it could be expected that sentence reading and reading comprehension explain more variance in teacher judgements than word reading. Additionally, self-concept likely determines how readily children engage in reading opportunities in front of their teachers, perhaps also affecting teacher assessments.

Rationale for the study

To summarise, there is strong evidence for the influence of reading fluency at the word and sentence level on reading comprehension in the English language. Nevertheless, evidence on this relationship in the orthographically regular language of German is scarce. Due to the greater consistency in grapheme-phoneme-correspondences, reading fluency should be attained more rapidly and consequently, there should be a pronounced correlation between sentence reading fluency and comprehension measures already at an earlier stage of reading and writing development in German.

Consequently, we address the following question: Does reading fluency at a word and sentence level have an incremental influence on reading comprehension at text level, going beyond pure grade level? We consequently expect higher order processes like sentence comprehension and reading fluency to become increasingly important already early on, while pure word reading accuracy should lose significance as children automate reading skills.

Further predictors likely explain variance above and beyond reading fluency and reading comprehension. First, there are clear indications of an association between AWM and reading fluency as well as reading comprehension (Cain et al., 2004). Because AWM is a predictor of reading comprehension independently from reading fluency, the question arises, if and to what extent it explains additional variance in reading comprehension beyond reading fluency. Thereby, we seek to test how the connection between reading fluency at word and sentence levels and reading comprehension is influenced by AWM. The same question applies for word recognition skills, as proposed in the non-lexical route of the dual route model (Coltheart et al., 2001). Because in transparent orthographies, direct lexical access quickly dominates the reading process (Rakhlin, Mourgues, Cardoso-Martins, Kornev, & Grigorenko, 2019), we assessed whether non-lexical recoding, measured by pseudoword reading, influences the connection between reading fluency on word and sentence level and reading comprehension.

Our third question addresses the assessment of reading skills by the teacher in these first stages of reading and spelling acquisition. There is currently scant evidence on how different predictors of reading comprehension influence the judgement of the teacher, and at the same time, these assessments play a vital role for the further academic career of the children in the first grades. We thus model teacher assessment as a function of different predictors (student's self-concept, pseudoword reading, fluency on word and sentence level, reading comprehension and AWM), to explore what influences assessments of reading fluency and reading comprehension. This last analysis is explorative in nature. It explores whether AWM, self-concept, and students' performances on reading fluency and reading comprehension predict teacher judgements of reading fluency and comprehension.

Methods

Materials

Reading skills were measured with the three subtests of the ELFE-II reading comprehension test (sight word reading, sentence fluency and passage comprehension) in the first study. In the second study, we administered measures of word and pseudoword decoding, sight word reading, sentence fluency and passage comprehension with the ELFE-II, SLS 2-9 and SLRT-II, working memory with the WISC-IV, teacher assessments and reading self-concept. All tests were conducted in the German language.

Reading comprehension test for the first to seventh grade-version II (ELFE-II). The ELFE-II reading comprehension test (Lenhard, Lenhard, & Schneider, 2017) is a standardised test of children's reading comprehension from the end of grade 1 to the beginning of grade 7. It includes basic reading skills as well as higher reading processes with a focus on reading comprehension. The subtest on sight word reading consists of a word picture matching task with 75 items (four alternatives with a picture) and a time limit of 3 minutes. Sentence fluency is measured with a multiple choice maze task (five alternatives) with 35 items and a time limit of 3 minutes. The passage level subtest is based on a model with the three dimensions genre (expository vs. narrative), coherence (local vs. global) and text basis (verbatim vs. gist) with 25 items and a maximum time of 7 minutes. The items vary in length from 202 to 876 letters and include a text, a question, and four alternatives. The results of the three subtests are normed independently and the single norm

scores eventually form a comprehensive total score. The test is highly reliable with r_{tt} = .98 for the word level, r_{tt} = .94 for the sentence level, r_{tt} = .89 for the text level and r_{tt} = .96 for the comprehensive score (odd-even-split). Retest reliability after 4 weeks is high, r_{tt} = .93, and shows high validity with respect to teacher assessments, r_{mean} = .70, and other standardised tests (Lenhard et al., 2017). In the first analysis, we use the subtest level data and in the second sample the total score of the three subtests to assess the overall reading comprehension. The norms are based on a continuous norming approach (W. Lenhard & Lenhard, 2021) and feature a month-by-month granularity.

Salzburg reading screening for grades 2-9 (SLS 2-9). The SLS 2-9 (Mayringer & Wimmer, 2014) was administered in the second sample in Grades 2 to 4 to measure reading fluency and comprehension. It consists of a sentence verification task with 100 sentences in ascending complexity and has a pronounced speed component. The sentences are judged true/false, with a time limit of 3 minutes. It thus combines reading accuracy as well as reading speed and has excellent reliability, $r_{tt} > .95$ in second grade, and validity, $r = .80$ to $r = .90$ with oral reading speed in second grade. The norms feature a year-by-year granularity. In Study 2, scores are reported as a reading quotient (M = 100, SD = 15) for sentence reading fluency.

Salzburg reading and spelling test (SLRT II). The SLRT II (Moll & Landerl, 2014) was used to differentially diagnose word reading skills with regard to the lexical and non-lexical processes in study 2. It contains a spelling section and two 1-minute word level decoding tests, in which as many words as possible are read aloud in 1 minute, followed by the same but with pseudowords. The decoding tests have a high reliability, $r_{tt} > .90$, and validity with respect to other reading tests. Raw scores represent the number of words read correctly.

Auditory working memory. Tests from the fourth edition of the Wechsler intelligence test for children (Wechsler, 2011) were used to measure AWM in study 2. The Wechsler intelligence scale is a highly reliable (r_{tt} = .97) and valid instrument. The subtest *Digit Span* is formed by *Digit Span Forward* and *Digit Span Backward*. The first test specifically measures the capacity of the phonological loop. Persons have to repeat orally presented number sequences from memory. The second test as well addresses executive functions. Again, orally presented numbers are repeated, but in reverse order. We applied the tests in an adapted, written form. The series of numbers were recited orally, but the children reproduced them on paper. There are norms for the age of 6 to 16 years with a quarterly granularity.

Reading self-concept questionnaire. In order to assess students' reading self-concepts, we used a self-created questionnaire based on the self-concept questionnaire from the KLASSE 4 performance inventory (Lenhard, Hasselhorn, & Schneider, 2011) in study 2. The questionnaire consisted of three items: (a) reading fluency and reading comprehension ('I can read texts quickly and understand them well'.), (b) performance compared with classmates ('Reading and understanding texts is much easier for me than for my classmates'.), and (c) grades in language arts ('I get good grades in language arts'). Ratings were collected on a 4-point visual-analogue Likert-scale, ranging from 1 = not true at all to 4 = completely true, in the form of smileys. In younger students below the age of 9 years,

the test showed a low homogeneity of ω = .52, but it increased with age to an acceptable homogeneity of ω = .67 for children aged 9 years and above.

Teacher assessment questionnaire. In order to measure teacher judgements' in study 2, teachers were asked to rate their students with regard to fluency and comprehension ('The student can read fluently' and 'The student can understand read content well') on a 4-point Likert-scale. Because the ratings were taken at the end of the school year, all participating teacher had an experience of at least one school year as the main teacher. A middle category was omitted in order to avoid a tendency towards the middle. The scores in both items ranged from 1 to 4 ($M_{Fluency}$ = 3.19, $SD_{Fluency}$ = 0.793, $M_{Comprehension}$ = 3.26, $SD_{Comprehension}$ = 0.702) and the intercorrelation and homogeneity amounted to r = .606 and ω = .76.

Procedure and sample description

Parents, teachers and schools gave their written consent for participation in the original studies. The data for the first research question were taken from the standardisation sample of ELFE-II (Lenhard et al., 2017) and hence represent secondary data and were collected in nine federal states in Germany from March to December 2015 in a group setting, totalling 5073 students from 71 schools. The testing per class took approx. 25 minutes. The representative norm sample was stratified according to region, school type, sex and family language background; 73.2% of the participants spoke exclusively German at home, 22.1% had a mixed family language, and 4.5% grew up monolingually in another language. ELFE-II includes two different versions (computer based versus paper based) with separable norms. The representative norm sample for the paper version includes 2056 first to sixth graders, and this dataset was used for the subsequent analysis. The age ranged between 5.5 and 14.5 years (M = 9.95 years, SD = 1.81 years; Table 1). The different grades included between n = 324 and n = 348 cases. Table 1 shows the demographic information and the number of students in the different school types for each grade level. Because the dataset represents a norm sample, the T scores for each subtest and the complete test approximated M = 50 and SD = 10 in the different grades and covered a range from $25 \leq T \leq 75$.

The data collection for the second study took place in a group format and was primary data. The students worked through the paper form of the tests in the following order: self-concept questionnaire, ELFE-II, SLS 2-9 and digit span from WISC-IV. The tests took 90 minutes. The SLRT II 1-minute reading fluency test was then conducted as a short individual test. The sample included 47 second, 28 third and 37 fourth graders (N = 112, 53.0% female) of a primary school in northern Bavaria. We collected the data in May 2019. The average age was M = 7.7 years (SD = 0.5 years) in second, M = 8.8 years (SD = 0.4 years) in third and M = 9.6 years (SD = 0.7 years) in fourth grade. In the overall sample, 74.3% of the children were monolingual German native speakers, 15.9% grew up with a mixed and 5.3% with a family language other than German, which is as well in line with national census data (Statistisches Bundesamt, 2015). ELFE-II can be regarded as a general measure of reading competence. Mean comprehension scores were slightly above average (Table 3). No child fell below the critical cut-off value of being strongly below average; however, one child was strongly above average. Due to illness, data from two children were missing for the word reading fluency; for the sentence reading fluency,

Table 1. Demographic variables for the first sample and intercorrelations of the subtests per scale in Study 1

Grade level	Age	Sex	School type								Correlations		
	M (*SD*)	% Female	Elem.	Comp.	Sec. M.	Jun. H.	Gram.	Spec. E.	Alt.	*N*	Word–Sentence	Sentence–Text	Word–Text
1	7.4 (0.4)	45.0	282	–	–	–	–	36	13	331	.768	.645	.566
2	8.4 (0.6)	48.8	306	–	–	–	–	20	18	344	.829	.763	.647
3	9.3 (0.7)	49.7	318	–	–	–	–	17	23	358	.808	.794	.642
4	10.3 (0.8)	51.8	330	–	–	–	–	13	14	357	.773	.777	.550
5	11.2 (0.6)	47.9	22	69	74	33	132	–	–	330	.754	.696	.553
6	12.2 (0.6)	42.3	33	57	68	61	119	–	–	338	.719	.651	.491

Notes: Elem. = Elementary; Comp. = Comprehensive; Sec. M. = Second. Modern; Jun.H. = Junior High; Gram. = Grammar; Spec.E. = Special Education; Alt. = Alternative school forms like Montessori or Waldorf. All correlations are significant on a level of $p < .001$.

one test had to be excluded due to incorrect completion. One questionnaire for the teacher's assessment was not returned, resulting in the drop out of one class with regard to teacher judgements.

Data analysis

For the first research question (i.e., word vs. sentence reading as a predictor of comprehension), we used the data from the first sample and included the word level test scores as a measure for word reading fluency, the sentence level test scores as a measure for sentence reading fluency, and the passage level test scores as a measure for reading comprehension. We then conducted hierarchical regression analyses predicting reading comprehension from word and sentence level measures to assess the moderating role of schooling duration. Because the testing was conducted in different waves and spread over 9 months of the school year, we included months of schooling as a more accurate control variable than grade in this study.

For the second and the third research questions, we used the data from the second sample. We included the measures of self-concept and teacher judgements, the test results of word reading fluency, pseudoword reading, and sentence reading fluency, the results of reading comprehension as the compound score of the subtests word, sentence, and text comprehension and the results for AWM as the compound score of the subtests Digit Span Forward and Backward. We then calculated several hierarchical regression analyses.

Results

Predicting reading comprehension by reading fluency

The first research question in this study is whether reading fluency at word and sentence level has an influence on reading comprehension that goes beyond pure age processes, and how this relationship changes over time. The descriptive raw scores of the variables from the first sample were, $M = 44.8$ and $SD = 17.58$ ($N = 2056$), for the word level test, $M = 19.5$ and $SD = 9.89$ ($N = 2056$) for the sentence level test, and $M = 12.4$ and $SD = 7.17$ ($N = 2055$) for the passage level test. The correlations between the variables were, $r = .912$, $p < .001$, for word and sentence level, $r = .794$, $p < .001$, for word and passage level, and, $r = .857$, $p < .001$, for sentence and passage level (see Table 1). A hierarchical regression analysis assessed the prediction of reading comprehension through word and sentence reading fluency as a function of grade (Table 2). Reading comprehension served as the criterion, while the schooling month was added as a predictor in the first step, reading fluency on word level in the second and reading fluency on sentence level in the third. In a fourth step, the interactions of schooling month with word and sentence reading fluency were added to assess if the interrelationship of these components changes over time.

All steps led to a significant increase in ΔR^2, with the final model reaching $R^2 = .743$, $F(5, 2040) = 1179.968$, $p < .001$ (Table 2). After inclusion of the interaction of sentence reading fluency and schooling month, word reading fluency and the interaction between word reading fluency and schooling duration lost their predictive power (word reading fluency, $\beta = .097$, $p = .138$; interaction schooling month and word reading fluency $\beta = -.100$, $p = .956$. The hypothesis that word and sentence reading fluency have an

Table 2. Hierarchical regression analysis predicting reading comprehension in Study 1

Steps and predictors	*b*	*SE b*	β	*t*	*p*
Step 1:					
Month of schooling	0.235	0.006	.674	41.24	<.001
R^2 = .454 ($p < .001$)					
Step 2:					
Month of schooling	0.064	0.007	.185	9.39	<.001
Word reading fluency	0.267	0.008	.657	33.31	<.001
ΔR^2 = .192 ($p < .001$)					
Step 3:					
Month of schooling	0.010	0.006	.030	1.64	.102
Word reading fluency	0.027	0.011	.066	2.34	.019
Sentence reading fluency	0.561	0.021	.775	26.37	<.001
ΔR^2 = .090 ($p < .001$)					
Step 4:					
Month of schooling	0.083	0.013	.238	6.35	<.001
Word reading fluency	0.040	0.027	.097	1.48	.138
Sentence reading fluency	0.652	0.049	.901	13.30	<.001
Month of schooling × Word reading fluency	0.000	0.001	−.100	−.87	.383
Month of schooling × Sentence reading fluency	−0.002	0.001	−.256	−2.47	<.013
ΔR^2 = .007 ($p < .001$)					
Overall-R^2 = .743 ($p < .001$)					

Note: N = 2046.

incremental influence on reading comprehension is partly confirmed. The significant interaction between sentence reading fluency and schooling month indicates a changing influence of the sentence reading fluency depending on the schooling. To further investigate this finding, we conducted correlation analyses for each grade between the variables word and sentence reading fluency and reading comprehension (Table 1).

Descriptively, the non-significant relationship between word reading fluency and reading comprehension in the final regression model displays a U-shaped development, beginning with a correlation of, r = .566, in first grade, reaching a peak in the second and third grade, r =. 647, and, r = .642, before gradually and significantly declining to, r = .491, in sixth grade, z = 3.056, $p < .001$. The correlation of sentence reading fluency and reading comprehension exhibits a similar, albeit stronger pattern, starting with, r = .645, in first grade, reaching its peak in third grade with, r = .794, and then significantly declining to, r = .651 in sixth grade, z = 4.001, $p < .001$. In sum, there is a developmental change in the relation of reading fluency and reading comprehension over the course of elementary grade with a peak between age 7 and 8 in second and third grade.

The role of the auditory working memory

Our next question addressed additional variance explained by AWM in predicting reading comprehension. We again performed a hierarchical regression analysis on sample 2

Table 3. Sample size, mean values and standard deviations of the variables from Study 2

	Descriptive data			Correlations						
	N	*M*	*SD*	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Word reading fluency	111	51.5	7.65							
(2) Pseudo word reading	111	51.2	7.87	.821						
(3) Sentence reading fluency	112	81.8	13.10	.542	.410					
(4) Reading comprehension	112	52.8	8.61	.453	.398	.515				
(5) Auditory working memory	113	57.3	11.50	.156^{+}	.105^{+}	.186*	.292**			
(6) Self-concept	113	3.06	0.47	.489	.481	.498	.491	.283**		
(7) Teacher judgement reading fluency	94	3.19	0.79	.434	.340	.378	.516	.447	.458	
(8) Teacher judgement reading comprehension	94	3.26	0.70	.344	.192^{+}	.409	.505	.254*	.461	.606

Notes: All descriptive data depicted as *T* scores (except from sentence reading fluency: $M = 100$, $SD = 15$; self-concept and teacher judgements: Likert scale 1–4). All unmarked correlations are significant on a level of $p = .001$; otherwise, $^{+}$n.s.
$^{*}p < .05$.
$^{**}p < .01$.

Table 4. Hierarchical regression analysis predicting reading comprehension in Study 2

Steps and predictors	*b*	*SE b*	β	*t*	*p*
Step 1:					
Auditory working memory	0.216	0.069	.295	3.19	.002
$R^2 = .087$ ($p = .002$)					
Step 2:					
Auditory working memory	0.145	0.059	.198	2.44	.016
Word reading fluency	0.259	0.107	.231	2.43	.017
Sentence reading fluency	0.224	0.062	.345	3.62	<.001
$\Delta R^2 = .250$ ($p < .001$)					
Overall-$R^2 = .337$ ($p < .001$)					

Note: $N = 110$.

(Table 3). Reading comprehension served as the criterion variable, to which predictor variables were added in two steps; first the AWM, then the two measures of reading fluency (word and sentence level; Table 4).

Both groups of predictors together explained an overall share of variance of 33.7%. The final model is as also significant, $F(3, 105) = 17.8$, $p < .001$. In the first step, the AWM proved to be a moderate predictor of reading comprehension, $\beta = .295$, $p = .002$, and its influence is maintained after adding reading fluency, $\beta = .198$, $p = .016$. Word, $\beta = .231$, $p = .017$, and sentence reading fluency, $\beta = .345$, $p < .001$, contributed significantly to the second model. Thus, AWM independently and in addition to fluency impacts reading comprehension.

Word recognition skills as a moderator

To answer the question of how pseudoword reading affect the relationship between reading fluency and reading comprehension, and how this relationship changes over time, we calculated a hierarchical regression in two blocks, separated by grade level. The results for the second grade (see Table 5) show that pseudoword reading is initially a significant predictor of reading comprehension, $\beta = .428$, $p = .004$; $R^2 = .183$, $p = .004$, but this predictive power disappears in the final model, $F(3, 40) = 9.80$, $p < .001$. Here, only sentence reading fluency, $\beta = .527$, $p < .001$, appears as a significant predictor. The overall model explains 42.4% of the variance of reading comprehension.

A similar pattern occurs in third grade; however, with pseudoword reading displaying only a trend in the first step, $\beta = .334$, $p = .077$. In the second step, again only the sentence reading fluency proves to be a valid predictor, $\beta = .517$, $p = .005$.

In fourth grade, pseudoword reading again reaches significance only in the first step, $\beta = .475$, $p = .003$. In the second step, none of the investigated predictors became significant. The second model does not explain incremental variance to the first model, $F(1, 34) = 9.9$, $p = .003$, which explains 22.6% of the variance. In sum, the basic word recognition skills, measured as accurate pseudoword reading, do not exert an influence on reading comprehension in any of the three grade levels, when reading fluency is statistically controlled.

Prediction of the teacher judgements

The third focus targeted the prediction of teacher judgements on reading fluency and reading comprehension. To this end, we again conducted two hierarchical regression analyses in two blocks with the dependent variables teacher judgement on reading fluency and teacher judgement on reading comprehension. We added the variables AWM and self-concept in the first and the variables word reading fluency, pseudoword reading, sentence reading fluency, and reading comprehension in the second block (Table 6). For the reading fluency judgement, the second model was statistically significant, $F(6, 89) = 10.2$, $p < .001$. AWM and the actual reading comprehension of the students were significant predictors of teacher's judgement on reading fluency, $\beta = .309$, $p < .001$, for AWM and, $\beta = .237$, $p < .030$, for reading comprehension. Self-concept, $\beta = .145$, $p < .001$, was significant only in the first model. The final model explained 41.8% of the total variance. Teacher judgement on reading comprehension were predicted by the students' self-concept, $\beta = .234$, $p = .041$, and reading comprehension, $\beta = .309$, $p < .008$. The final model accounted for 34.8% of the variance.

Table 5. Hierarchical regression analysis predicting reading comprehension in grades 2 to 4 in Study 2

	Grade 2 (N = 47)					Grade 3 (N = 29)					Grade 4 (N = 36)				
Steps and predictors	*b*	*SE b*	β	*t*	*p*	*b*	*SE b*	β	*t*	*p*	*b*	*SE b*	β	*t*	*p*
Step 1:															
Pseudo word reading	.405	.132	.428	3.07	.004	.403	.219	.334	1.84	.077	.554	.176	.475	3.15	.003
	R^2 = .183 (p = .004)					R^2 = .111 (p = .077)					R^2 = .226 (p = .003)				
Step 2:															
Pseudo word reading	.075	.221	.080	.34	.735	−.192	.272	−.159	−0.71	.487	.384	.319	.330	1.21	.237
Word reading fluency	.135	.246	.131	.55	.587	.566	.368	.379	1.54	.137	.086	.329	.084	.26	.795
Sentence reading fluency	.327	.086	.527	3.82	< .001	.397	.130	.517	3.05	.005	0.126	.119	.208	1.06	.297
	ΔR^2 = .240 (p < .001)					ΔR^2 = .379 (p < .001)					ΔR^2 = .050 (p = .343)				
	Overall-R^2 = .424 (p < .001)					Overall-R^2 = .490 (p < .001)					Overall-R^2 = .276 (p = .015)				

Table 6. Hierarchical regression analysis predicting the teacher judgement on reading fluency and reading comprehension in Study 2

	Reading fluency					Reading comprehension				
Steps and predictors	*b*	*SE b*	β	*t*	*p*	*b*	*SE b*	β	*t*	*p*
Step 1:										
Auditory working memory	0.026	0.007	.346	3.84	<.001	0.009	0.006	.140	1.44	.154
Self-concept	0.613	0.158	.145	3.88	<.001	0.634	0.149	.413	4.25	<.001
	R^2 = .316 (p < .001)									
Step 2:										
Auditory working memory	0.023	0.007	.309	3.47	<.001	0.004	0.006	.056	0.59	.558
Self-concept	0.231	0.184	.133	1.25	.213	0.359	0.173	.234	2.08	.041
Word reading fluency	0.019	0.017	.177	1.11	.271	0.021	0.016	.227	1.35	.182
Pseudo word reading	0.001	0.016	.010	0.07	.947	−0.025	0.015	−.269	−1.73	.087
Sentence reading fluency	0.003	0.006	.051	0.47	.637	0.006	0.006	.110	.973	.333
Reading comprehension	0.022	0.010	.237	2.20	.030	0.026	0.009	.309	2.72	.008
	ΔR^2 = .102 (p = .008)					ΔR^2 = .126 (p = .004)				
	Overall-R^2 = .418 (p < .001)					Overall-R^2 = .348 (p < .001)				

Note: N = 95.

Discussion

The goal of this study was, broadly, to investigate the determinants of and the relationship between reading fluency and reading comprehension in a more regular orthography, namely, in German, and how these determinants influence the formation of teacher judgements. We first investigated whether reading fluency played a role in children's reading comprehension performance, depending on whether this was measured with word or sentence reading and, secondly, how this changed across the grades. Results indicated that sentence reading fluency was the strongest predictor of reading comprehension for passages. Further, there appeared to be a clear developmental pattern, as hypothesized, with word reading fluency losing its predictive power across the grades. A closer look at sentence reading fluency shows that the strength of this influence seems to change over the course of development, increasing in the first grades, reaching a peak in the third grade and slowly decreasing until the sixth grade. This finding supports the results, for example, of Kim et al. (2012), who found a decreasing significance of word reading fluency and an increasing significance of sentence reading fluency for children with increasing reading experience and it is in line with Landerl, Wimmer, and Frith (1997) that children reading in German gain fluency much earlier.

Our assumption that the significance of reading fluency at first should increase rapidly in an orthographically consistent language such as German therefore was confirmed. Although the strong association between sentence reading fluency and reading comprehension decreases after a peak around the third grade, the relationship of both concepts stays at a moderate level until sixth grade. This decrease could be due to the increasing importance of higher-order processes and other determinants such as vocabulary or prior knowledge. Interestingly, our results do not support the conclusions of Schwanenflugel et al. (2006), who do not regard sentence reading fluency as a relevant predictor of reading comprehension in the lower grades. Thus, the current study also contributes to research from similarly regular orthographies (e.g., Kim et al., 2012; Kim, Park, & Wagner, 2014), together indicating a faster development of sight word reading and fluency compared with less regular orthographies (e.g., Seymour et al., 2003). When directly comparing the developmental speed in decoding words and pseudowords between English and German, (Frith, Wimmer, & Landerl, 1998; Landerl et al., 1997), English speaking children from grade 1 to 3 need more than double the time in decoding and they make more three times more error than their German age and grade level peers, even when the item material is highly parallelized. Accordingly, we speculate that pseudoword decoding loses its predictive relevance in more regular languages.

Turning to the second aim of investigating the role of the AWM, our results show that it makes a significant contribution to the variance explained in reading comprehension, even when reading fluency is controlled for. In other words, the better the students were able to store information in working memory, the better their performance in reading comprehension. Previous studies (Cain et al., 2004; Wagner et al., 1994) have found a moderate to high correlation between AWM and reading comprehension, and yet there are also opposing views to this position (e.g., Hutton & Towse, 2001). In the sense of the simple view of reading (Gough & Tunmer, 1986), AWM cannot only be attributed to phonological information processing skills, but probably also to linguistic skills and construction of situational models (Kintsch, 1998). However, this finding needs to be tempered by our not measuring phonological memory. Specifically, within the framework of phonological information processing (Wagner & Torgesen, 1987), AWM and

phonological awareness share common variance and hence either might drive the correlation (e.g., Siu et al., 2018).

We also examined the contribution of word recognition skills as proposed in the indirect route of the dual route model by Coltheart et al. (2001), measured by accurate pseudoword decoding. Findings indicate that these did not make a significant contribution to reading comprehension in any of the three grade levels, when word and sentence reading fluency were taken into account. One possible explanation would be that in a shallow orthography such as German, the transition to automated reading occurs very early on, possibly already at the end of first grade. Therefore, the influence of the indirect route decreases more rapidly than in English, where this change occurs at about the age of 10 (García & Cain, 2014). However, to our knowledge, there are no studies on whether reading accuracy predicts reading comprehension beyond reading fluency in the first grade in German, investigating this question would be a possible subject for future research to better understand the transition from non-lexical to lexical reading.

The last research question focused on determinants of the very brief and subjective teacher assessments concerning the overall reading fluency and reading comprehension. AWM and the actual reading comprehension performance were the best predictors of teachers' reading fluency judgements, whereas the actual reading fluency performance seemed to be statistically less significant. The best predictors of reading comprehension judgements were the variables self-concept and actual reading comprehension performance. The fact that students' comprehension performance is a significant predictor of both teacher assessments actually supports the validity of the teacher judgement with regard to the relative position of the children, especially with respect to the reading comprehension assessment of the teachers, but it as well shows the limitation of subjective teacher assessments. Especially because teachers only poorly differentiated between fluency, comprehension, self-concept, and AWM, they might draw the wrong conclusions and subsequently select inappropriate instructional or remedial measures, or even make incorrect decisions about children's future school careers. Standardised tests could thus help support judgements, better tailor pedagogical decisions and interventions to children's ability levels and ensure that children with poor performance but a strong self-concept are not overlooked.

Limitations and future research

We used a fairly brief overall teacher judgement assessment, consisting of only two items, ranging from 1 to 4. The same applies for the measure of self-concept, which relies on only few items and consequently displays a low homogeneity. Further, it would have been insightful, although practically difficult, to determine how teachers formed their assessment, possibly through interviews. Additionally, our study is not directly comparable with work conducted in other orthographies (e.g., e.g., Hoover & Gough, 1990; Kim et al., 2012; Schwanenflugel et al., 2006) due to our using different measures. Accordingly, we cannot be certain that the findings would transfer to similarly transparent writing systems. Future research should employ cross-language samples using measures that are matched as closely as possible.

Interestingly, students' self-concepts seemed to be a moderate predictor of teacher judgements of reading comprehension. The students' self-assessments consisted of one variable for self-assessment with respect to their own abilities in the area of reading fluency and reading comprehension, one for self-assessment in social comparison and one for

self-assessment of their German grades. On one hand, a correspondence between perceived and objective assessment by the teacher again speaks for the validity of the teacher's judgement. On the other hand, it would be an interesting question for future research as to whether and what extent there is a possible Rosenthal effect, in which the teacher's assessment influences the students' assessment and thus their self-concept, which in turn influences the teacher's assessment.

Data Availability Statement

The data of the elementary school sample are freely available via https://osf.io/3c48t/. The normative sample data of ELFE II will be provided on request to the second author, who is as well co-author of this reading comprehension test.

References

Adams, M.J. (1990). *Beginning to read: Thinking and learning about print*. The MIT Press.

Statistisches Bundesamt (2015). Schüler: Bundesländer, Schuljahr, Geschlecht, Jahrgangsstufen, Schulart [Students: Federal states, school year, sex, grade, school type]. Retrieved from https://www-genesis.destatis.de/genesis/online/data

Cain, K., Oakhill, J. & Bryant, P. (2004). Children's reading comprehension ability: Concurrent predicition by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, 96(1), 31–42. https://doi.org/10.1037/0022-0663.96.1.31

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

Cromley, J.G. & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, 99(2), 311–325. https://doi.org/10.3102/0034654313499616

Ehri, L. C. (2005). Development of sight word reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading*. Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch8

Ennemoser, M., Marx, P., Weber, J. & Schneider, W. (2012). Spezifische Vorläuferfertigkeiten der Lesegeschwindigkeit, des Leseverständnisses und des Rechtschreibens [Specific precursor skills of reading speed, reading comprehension, and spelling]. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, 44(2), 53–67. https://doi.org/10.1026/0049-8637/a000057

Feinberg, A.B. & Shapiro, E.S. (2003). Accuracy of teacher judgments in predicting oral reading fluency. *School Psychology Quarterly*, 18(1), 52–65. https://doi.org/10.1521/scpq.18.1.52.20876

Feinberg, A.B. & Shapiro, E.S. (2009). Teacher accuracy: An examination of teacher-based judgments of Students' Reading with differing achievement levels. *The Journal of Educational Research*, 102(6), 453–462. https://doi.org/10.3200/JOER.102.6.453-462

Frith, U., Wimmer, H. & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading*, 2(1), 31–54. https://doi.org/10.1207/s1532799xssr0201_2

García, J.R. & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, 84(1), 74–111. https://doi.org/10.3102/0034654313499616

Gough, P.B. & Tunmer, W.E. (1986). Decoding, reading and reading disability. *Remedial & Special Education*, 7, 6–10.

Hattie J. (2012). *Visible learning for teachers: Maximizing impact on learning*. London: Routledge.

Herman, J.A., Matyas, T. & Pratt, C. (2006). Meta-analysis of the nonword reading deficit in specific reading disorder. *Dyslexia*, 12(3), 195–221. https://doi.org/10.1002/dys.324

Hoge, R.D. & Coladarci, T. (1989). Teacher-based judgments of academic achievement. *Review of Educational Research*, 59(3), 297–781. https://doi.org/10.3102/00346543059003297

Hoover, W.A. & Gough, P.B. (1990). The simple view of reading. *Reading and Writing*, 2(2), 127–160. https://doi.org/10.1007/BF00401799

Hutton, U.M. & Towse, J.N. (2001). Short-term memory and working memory as indices of children's cognitive skills. *Memory*, 9(4-6), 383–394. https://doi.org/10.1080/09658210042000058

Jenkins, J.R., Fuchs, L.S., van den Broek, P., Espin, C. & Deno, S.L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, 95(4), 719–729. https://doi.org/10.1037/0022-0663.95.4.719

Joshi, R.M., Tao, S., Aaron, P.G. & Quiroz, B. (2012). Cognitive component of componential model of reading applied to different orthographies. *Journal of Learning Disabilities*, 45(5), 480–486. https://doi.org/10.1177/0022219411432690

Kaiser, J., Südkamp, A. & Möller, J. (2017). The effects of student characteristics on teachers' judgment accuracy. *Journal of Educational Psychology*, 109(6), 871–888. https://doi.org/10.1037/edu0000156

Karing, C., Matthäi, J. & Artelt, C. (2011). Genauigkeit von Lehrerurteilen über die Lesekompetenz ihrer Schülerinnen und Schüler in der Sekundarstufe I [accuracy of teacher judgements about the literacy of their students in lower secondary education]. *Zeitschrift für Pädagogische Psychologie*, 25(3), 159–172. https://doi.org/10.1024/1010-0652/a000041

Keenan, J.M., Betjemann, R.S. & Olson, R.K. (2008). Reading comprehension tests vary in the skills they assess. *Scientific Studies of Reading*, 12(3), 281–300. https://doi.org/10.1080/10888430802132279

Kibby, M.Y., Lee, S.E. & Dyer, S.M. (2014). Reading performance is predicted by more than phonological processing. *Frontiers in Psychology*, 5, 960. https://doi.org/10.3389/fpsyg.2014.00960

Kim, Y., Wagner, R.K. & Lopez, D. (2012). Developmental relations between reading fluency and reading comprehension. *Journal of Experimental Child Psychology*, 113(1), 93–111. https://doi.org/10.1016/j.jecp.2012.03.002

Kim, Y.G., Park, C.H. & Wagner, R.K. (2014). Is oral/text reading fluency a "bridge" to reading comprehension? *Reading & Writing*, 27(1), 79–99. https://doi.org/10.1007/s11145-013-9434-7

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge university press.

Klauda, S.L. & Guthrie, J.T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, 100(2), 310–321. https://doi.org/10.1037/0022-0663.100.2.310

Kriegbaum, K., Steinmayr, R. & Spinath, B. (2019). Longitudinal reciprocal effects between teachers' judgments of students' aptitude, students' motivation, and grades in math. *Contemporary Educational Psychology*, 59, 101807. https://doi.org/10.1016/j.cedpsych.2019.101807

Landerl, K., Wimmer, H. & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German-English comparison. *Cognition*, 63(3), 315–334. https://doi.org/10.1016/S0010-0277(97)00005-X

Leather, C.V. & Henry, L.A. (1994). Working memory span and phonological awareness tasks as predictors of early reading ability. *Journal of Experimental Child Psychology*, 58(1), 88–111. https://doi.org/10.1006/jecp.1994.1027

Lenhard, W., Hasselhorn, M. & Schneider, W. (2011). *Kombiniertes Leistungsinventar zur allgemeinen Schulleistung und für Schullaufbahnempfehlungen in der vierten Klasse [combined performance inventory for general school performance and for school career recommendations in the fourth grade]*. Hogrefe.

Lenhard, W. & Lenhard, A. (2021). Improvement of norm score quality via regression-based continuous norming. *Educational and Psychological Measurement*, 81(2), 229–261. https://doi.org/10.1177/0013164420928457

Lenhard, W., Lenhard, A. & Schneider, W. (2017). *Ein Leseverständnistest für Erst- bis Siebtklässler (ELFE-II) [a reading comprehension test for students from grade 1 to 7]*. Hogrefe.

Leong, C.K., Tse, S.K. & Hau, K.T. (2008). Text comprehension in Chinese children: Relative contribution of verbal working memory, pseudoword reading, rapid automatized naming, and onset-rime phonological segmentation. *Journal of Educational Psychology*, 100(1), 135–149. https://doi.org/10.1037/0022-0663.100.1.135

Mayringer, H. & Wimmer, H. (2014). *Salzburger Lesescreening für die Schulstufen 2–9 (SLS 2–9) [Salzburg reading screening for Grade 2 to 9]*. Hogrefe.

Moll, K. & Landerl, K. (2014). *Salzburger Lese- und Rechtschreibtest II [Salzburg test on reading and spelling]*. Hogrefe.

Näslund, J.C. (1990). The interrelationships among preschool predictors of reading acquisition for German children. *Reading and Writing*, 2(4), 327–360. https://doi.org/10.1007/BF00419387

Rakhlin, N.V., Mourgues, C., Cardoso-Martins, C., Kornev, A.N. & Grigorenko, E.L. (2019). Orthographic processing is a key predictor of reading fluency in good and poor readers in a transparent orthography. *Contemporary Educational Psychology*, 56, 250–261. https://doi.org/10.1016/j.cedpsych.2018.12.002

Rausch, T. (2016). Lehrerurteile über Schülerleistungen [Teacher assessments of student performance]. *Bamberg graduate school of social sciences*. Retrieved from https://opus4.kobv.de/opus4-bamberg/files/49558/RauschTobiasDissopusse_A3a.pdf

Schwanenflugel, P.J., Meisinger, E.B., Wisenbaker, J.M., Kuhn, M.R., Strauss, G.P. & Morris, R.D. (2006). Becoming a fluent and automatic reader in the early elementary school years. *Reading Research Quarterly*, 41(4), 496–522. https://doi.org/10.1598/RRQ.41.4.4

Segerer, R., Niklas, F., Suggate, S. & Schneider, W. (2020). Young minority home-language students' biased reading self-concept and its consequences for reading development. *Reading Research Quarterly*, 45(1), 31–94. https://doi.org/10.1002/rrq.300

Seigneuric, A. & Ehrlich, M. (2005). Contribution of working memory capacity to children's reading comprehension. *Reading and Writing*, 18(7-9), 617–656. https://doi.org/10.1007/s11145-005-2038-0

Seymour, P.H.K., Aro, M. & Erskine, J.M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143–174. https://doi.org/10.1348/000712603321661859

Siu, T.C., McBride, C., Tse, C., Tong, X. & Maurer, U. (2018). Evaluating the effects of metalinguistic and working memory training on reading fluency in Chinese and English. *Frontiers in Psychology*, 9, 2510. https://doi.org/10.3389/fpsyg.2018.02510

Südkamp, A., Kaiser, J. & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement. *Journal of Educational Psychology*, 104(3), 743–762. https://doi.org/10.1037/a0027627

Suggate, S.P., Reese, E., Lenhard, W. & Schneider, W. (2014). The relative contributions of vocabulary, decoding, and phonemic awareness to word reading in English versus German. *Reading & Writing*, 27(8), 1395–1412. https://doi.org/10.1007/s11145-014-9498-z

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

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of word reading efficiency—second edition*. Pro-Ed.

van Kraayenoord, C.E. & Schneider, W.E. (1999). Reading achievement, metacognition, reading self-concept and interest. *European Journal of Psychology of Education*, 14(3), 305–324. https://doi.org/10.1007/BF03173117

von Goldammer, A., Mähler, C., Bockmann, A.-K. & Hasselhorn, M. (2010). Vorhersage früher Schriftsprachleistungen aus vorschulischen Kompetenzen der Sprache und der phonologischen Informationsverarbeitung [predicting early literacy performance from preschool competencies of language and phonological information processing]. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, 42(1), 48–56. https://doi.org/10.1026/0049-8637/a000005

Wagner, R.K. & Torgesen, J.K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101(2), 192–212. https://doi.org/10.1037/0033-2909.101.2.192

Wagner, R.K., Torgesen, J.K. & Rashotte, C.A. (1994). Development of reading related phonological processing abilities. *Developmental Psychology*, 30(1), 73–87. https://doi.org/10.1037/0012-1649.30.1.73

Wechsler, D. (2011). *Wechsler intelligence scale for children—Fourth edition (WISC IV)*. Pearson Assessment.

Zhu, M., Urhahne, D. & Rubie-Davies, C.M. (2018). The longitudinal effects of teacher judgement and different teacher treatment on students' academic outcomes. *Educational Psychology*, 38(5), 648–668. https://doi.org/10.1080/01443410.2017.1412399

Nicole Kirschmann finished her psychology studies at the University of Würzburg. She currently works on the field of learning disorders and is especially interested in development of reading skills and reading comprehension.

Wolfgang Lenhard is an associate professor, working in educational psychology at the University of Würzburg. His research interests are diagnosis and interventions in the field of reading comprehension, learning disorders, computer based assessment, intelligent tutorial systems, fostering mathematical abilities, cognitive trainings and diagnosis of attention-deficit hyperactivity disorder.

Sebastian Suggate is a senior lecturer at the University of Regensburg and conducts research into reading and language development and the contribution of fine and sensorimotor skills to learning and cognition.

Received 10 November 2020; revised version received 4 June 2021.

Address for correspondence: Wolfgang Lenhard, Department of Psychology, University of Wuerzburg, Wittelsbacherplatz 1, 97074 Wuerzburg, Germany. E-mail: wolfgang.lenhard@uni-wuerzburg.de