Secondary Publication



Tiffin-Richards, Simon P.; Lenhart, Jan; Marx, Peter

Children's metacognitive awareness and self-correction of their own spelling mistakes

Date of secondary publication: 28.09.2026

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-117371x

Primary publication

Tiffin-Richards, Simon P.; Lenhart, Jan; Marx, Peter (2026): Children's metacognitive awareness and self-correction of their own spelling mistakes, in: Metacognition and learning, Berlin, Heidelberg [u.a.]: Springer, Vol. 21, No. 1, pp. 1–21, doi: 10.1007/s11409-026-09491-z

Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available under a Creative Commons license.



The license information is available online:
https://creativecommons.org/licenses/by/4.0/legalcode

Metacognition and Learning (2026) 21:41
https://doi.org/10.1007/s11409-026-09491-z

Children's metacognitive awareness and self-correction of their own spelling mistakes

Simon P. Tiffin-Richards[1] · **Jan Lenhart**[2] · **Peter Marx**[3]

Received: 12 November 2025 / Accepted: 6 August 2026
© The Author(s) 2026

Abstract
This paper examines how children identify and correct spelling errors in their own writing. Our research focused on two questions: (a) how accurate primary school children are in their metacognitive monitoring of their own writing of a dictated text, and (b) whether they are able to translate their spelling performance monitoring into regulatory behaviour such that they revise and correct detected spelling errors. A total of 318 German-speaking primary school children wrote a continuous text that was dictated to them by an instructor in their language of instruction (German). In order to differentiate between the processes of monitoring and control, children were first instructed to underline words they were unsure about (Review Phase). After this initial monitoring process, children were told that they could change their initial spellings (Revise Phase). Our results demonstrate that the primary school children were accurate in identifying their correctly spelled words (high specificity), whereas their error detection rate was poor (low sensitivity). Furthermore, only about 24% of the words selected in the Review Phase were then actually changed in the Revise Phase. When children did revise the spelling of a word, they were more likely to correct an initial spelling mistake than change a correct spelling. Children's difficulties in identifying and correcting their own spelling mistakes seem to be related to their low sensitivity to spelling errors (spelling awareness deficit) and their reluctance or inability to correct misspelled words (implementation or knowledge deficit).

Keywords Metacognition · Spelling consciousness · Spelling · Self-correction · Answer changing

The ability to judge one's own performance, identify mistakes, and correct them requires metacognition: the ability to monitor and regulate one's own mental processes (Flavell,

✉ Simon P. Tiffin-Richards
tiffin-richards@leibniz-ipn.de

[1] Department of Educational Research and Educational Psychology, Leibniz Institute for Science and Mathematics Education, Kiel, Germany

[2] Department of Psychology, University of Bamberg, Bamberg, Germany

[3] Department of Psychology IV, University of Würzburg, Würzburg, Germany

Published online: 19 August 2026

1979). Monitoring one's thoughts and cognitive processes is an important factor in academic performance (Nietfeld et al., 2005; De Boer et al., 2018; Donker et al., 2014; Zohar & Barzilai, 2013), problem-solving (Georghiades, 2004), decision-making (Yeung & Summerfield, 2012), and even extends to emotion regulation (Pennequin et al., 2020). Metacognitive monitoring facilitates resource allocation such as study time (Son & Metcalfe, 2000), strategy selection, and error detection (Nelson, 1990), and is thus critical for self-regulated learning. Despite its wide-ranging relevance for children's development (Kuhn, 2000) and extensive research on metacognition in educational settings (see Baker, 1989, for a review on reading comprehension; see Schneider & Artelt, 2010, for a review on mathematics), we know surprisingly little about how metacognitive skills apply to writing and spelling.

Awareness of one's performance is clearly relevant to detecting one's own errors and addressing them in a constructive manner (Tulis et al., 2016; Zhang & Fiorella, 2023). In assessment contexts, students use their confidence as a cue to identify and revise incorrect answers on exam papers (Couchman et al., 2016; Stylianou-Georgiou & Papanastasiou, 2017; Tiffin-Richards et al., 2022). However, accurate self-assessment is not always successfully translated into self-regulatory behaviour. For instance, while children can accurately judge what they know (e.g., with judgements of learning, or judgements of knowing), they often fail to use their self-awareness to select appropriate study materials to facilitate the learning process (Metcalfe & Finn, 2013). This implies that children are susceptible to an *implementation deficit* (Metcalfe, 2009), whereby they sometimes fail to translate accurate monitoring of their performance into efficient behavioural strategies and regulation (Dufresne & Kobasigawa, 1989; Koriat et al., 2009), even when they have the necessary knowledge and strategies at their disposal. This paper examines children's capacity to identify and self-correct spelling errors in their own writing, also referred to as *spelling consciousness* (Lull, 1917), which may offer a window into processes of metacognitive monitoring and control in early literacy acquisition.

Metacognitive monitoring and control

Metacognition is commonly defined as "knowledge and cognition about cognitive phenomena" (Flavell, 1979, p. 906) that allows people to monitor their own memory, comprehension, and other cognitive processes. To understand why children may struggle to detect and correct errors in their writing, it is helpful to consider Nelson's (1990) framework of metacognition. This model distinguishes between *object-level* processes (e.g., encoding, learning, and remembering) and a higher-order *meta-level* that monitors and regulates these processes. *Metacognitive monitoring* refers to the awareness and evaluation of one's cognitive performance, while *metacognitive control* involves using this information to guide behaviour toward a specific goal. These two processes enable the meta-level to update its task model and adjust object-level activities accordingly—for example, recognizing a spelling error and deciding to attempt to correct it.

Measures of metacognitive monitoring Metacognitive monitoring is typically measured by collecting item-by-item ratings of participants' response confidence (Fleming & Lau, 2014; Schraw, 2009) or judgements of learning (Koriat, 2008), or with more general survey questionnaires of a person's metacognitive awareness (Schraw & Dennison, 1994). In the context

of item-by-item metacognitive monitoring, *calibration* refers to the relationship between task performance and a judgement about that performance (Boekaerts & Rozendaal, 2010). A wide range of statistical methods have been implemented to specify the calibration of item-by-item judgements and performance (extensively reviewed in Schraw et al., 2013 and Schraw et al., 2014). Hierarchical modelling conducted by Schraw et al. (2014) suggested that different measures may be best grouped together as measures of *diagnostic efficiency* (Benjamin & Diaz, 2008), including the accuracy of judgments about correct performance (i.e., *sensitivity*) and incorrect performance (i.e., *specificity*) that are subsumed under a measure of differentiation (Kadlec, 1999) between correct and incorrect performance (i.e., *discrimination*). In the context of spelling awareness, sensitivity can refer to the proportion of detected spelling errors and represents the accuracy of judgments about words containing spelling errors. Specificity, on the other hand, can represent the proportion of correct spellings that are not mistakenly judged as errors and corresponds to the accuracy of judgments about correctly spelled words. Put simply, high sensitivity means that a writer can detect most of their spelling errors, while high specificity means that a writer can recognize most of the words they have spelled correctly. A discrimination measure is a further index that describes whether an individual can distinguish between correct and incorrect performance (Schraw, 2009), in this case differentiating between a writer's own correctly and incorrectly spelled words.

Signal detection theory (SDT; Green & Swets, 1966; Swets, 1988) is a widely used methodological approach to characterize how sensitive a person is to a specific cue (Lynn & Barrett, 2014). For binary decisions (e.g., whether a cue, such as a spelling mistake, is present or not), decisions can be allocated to one of four categories: cue present and detected (Hit), cue present but undetected (Miss), cue not present but detected (False Alarm), and cue not present and not detected (Correct Rejection). In the present case, the cue represents an incorrect spelling pattern and the absence of the cue indicates that a word is spelled correctly. Children's discrimination between correct and incorrect spellings in their own writing can be conceptualized in the terms of SDT as the difference between correct detections of spelling mistakes (Hits) and incorrect detections of errors (False Alarms). The discrimination index d' is calculated as:

$$d' = z(H) - z(FA),$$

where H represents the proportion of hits (e.g., proportion of incorrect spellings judged to be errors) and FA represents the proportion of false alarms (e.g., proportion of correct spellings judged to be errors). Both distributions are z-standardized, meaning that the proportion of 0.5 is converted to a z-score of 0. In this context, larger absolute values of d' indicate that a person is more aware of whether a spelling is correct or not, while d' values near zero indicate chance performance.

Translating monitoring to control The detection and correction of errors is a process in which metacognition is critical (Yeung & Summerfield, 2012) and where *performance confidence* plays an important role (Fleming, 2024). Ratings of response confidence provided by examinees, which represent a common method for assessing online metacognitive monitoring (Lingel et al., 2019), correlate positively with response accuracy (Koriat,

2012; Nietfeld et al., 2005; Shaughnessy, 1979). This suggests that people usually have some understanding of their ability to solve certain problems or carry out specific tasks. In addition, it has been demonstrated that examinees' confidence in their initial response to a problem or question is negatively correlated with the probability of changing that response (Couchman et al., 2016; Stylianou-Georgiou & Papanastasiou, 2017; Tiffin-Richards et al., 2022). Response confidence hence appears to direct revision behaviour. Put another way, the process of *monitoring* confidence in initial answers appears to *regulate* examinees' answer changing behaviour.

Empirical studies of metacognitive monitoring and answer-changing (e.g., Cornoldi, 1990; Metcalfe & Finn, 2013) suggest that three challenges must be overcome for error revision to be successful: (1) accurate performance monitoring (metacognitive monitoring); (2) translation of monitoring judgements into regulatory behaviour, such as answer revision (metacognitive control); and (3) the domain-specific or procedural knowledge necessary to retrieve the correct answer from memory or solve the problem at hand (knowledge availability).

Spelling awareness

In the present study, we were interested in children's metacognitive monitoring and control processes in a domain specifically relevant for primary school children, namely their orthographic knowledge (Landerl et al., 2022) and spelling proficiency (Pan et al., 2021). The importance of metacognitive spelling awareness, also referred to as spelling consciousness, has been debated and studied in educational contexts for over a century (e.g., Lull, 1917; Pryor & Pittman, 1921). Broadly speaking, spelling consciousness is the ability to judge whether one's spelling is correct, or put more concretely, the "ability to know that a word one has written is spelled either correctly or incorrectly or that one is doubtful of the spelling." (Valmont, 1972, p. 1219), as well as the ability and predisposition to actively engage with one's own spelling errors (Pryor & Pittman, 1921). This definition thus appears to incorporate both monitoring (i.e., error detection) and control (i.e., error revision) processes.

Early studies of spelling consciousness and error detection (e.g., Grothe, 1966; McFarland, 1916; Tidyman, 1924) raised concerns about students' ability to find and correct spelling errors in their own writing. Their findings suggest that spelling consciousness is strongly related to spelling ability, as well as to general intelligence, and is variable within individuals and across contexts (Valmont, 1972). Being aware of one's own spelling errors appears to be positively related to spelling performance (Block & Peskowitz, 1990) and to the selection of appropriate spelling strategies during writing (Kreiner & Green, 2000).

Intervention studies have found that training students' ability to self-correct their own spelling can improve overall spelling ability in primary school children (McGuffin et al., 1997; Silva et al., 2021). Spelling consciousness, measured as the relation between students' confidence in their spelling and their spelling accuracy, has also been shown to improve with explicit training in college students (Kreiner & Green, 2000).

Interventions specifically aimed at improving children's metacognitive spelling awareness have also produced positive short-term effects. Cordewener et al. (2016) compared three training conditions in a between-groups design with 72 third-grade Dutch-speaking students. All students completed four dictation sessions. In each session, they first judged

whether they could spell each word before writing it (i.e., pre-writing metacognitive judgements). After each session, they received different training: those in the *strategy-instruction* group divided words into syllables and applied spelling rules to correct them; those in the *self-correction* group used an answer sheet to correct their errors; and the *no-correction* group received no feedback or further practice. The short interventions improved both spelling and awareness compared to the control group, but effects were not maintained at a five-week follow-up. Strategy instruction reduced response bias (over-confidence in their own spelling), measured using signal detection theory, suggesting more accurate self-judgments.

In a follow-up study, Cordewener et al. (2018) compared four training conditions with 115 third-grade Dutch-speaking students. In the *strategy-instruction* condition, students were guided to think about how to spell a word before writing it, using syllable-based segmentation. In the *strategic-monitoring* condition, students reflected on their spelling after writing dictated words, correcting their mistakes using an answer sheet and syllable-based segmentation. An instructor asked additional questions about their thinking to stimulate children's metacognitive processes. In the *self-monitoring* condition, students checked their own spellings using an answer sheet without further instruction. The *no-correction* group received no feedback or further practice. Spelling consciousness was measured by asking students, before writing, to indicate whether they thought they could spell each word correctly by circling *yes* or *no* on their sheet. Their results found that all training conditions outperformed the control group in a post-training spelling test, while only the strategy-instruction group improved in their metacognitive judgement accuracy. Cordewener et al. (2018) did not include a follow-up assessment of maintained training effects.

These intervention studies suggest that metacognitive spelling awareness can be trained using direct strategy instruction and practice with positive effects on spelling ability, although evidence of lasting benefits is lacking. Furthermore, it is unclear to what extent error detection leads to error correction in students' own writing, as previous studies have mainly focused on error awareness (metacognitive monitoring; e.g. Cordewener et al., 2016; Cordewener et al., 2018), or error correction (metacognitive control; e.g. Silva et al., 2021). There has been little research into the relationship between these two metacognitive processes in the context of writing and spelling. It also remains unclear to what extent individual differences in factors such as language ability, age, and years of formal literacy education influence metacognitive monitoring and control when correcting one's own writing.

A developmental perspective

Studying spelling awareness and self-correction is particularly interesting in primary school children at an age where they are in principle able to metacognitively monitor and control their behaviour and have the orthographic knowledge necessary to be able to detect and correct spelling errors. There is converging evidence that metacognitive monitoring and control emerge early and improve with age across childhood, though not necessarily at the same pace. Studies conducted with 8 to 12-year-olds show that children can distinguish between easy, difficult, and unanswerable questions, differentiate correct from incorrect answers via confidence judgments (monitoring), and selectively withdraw likely incorrect responses (control), with monitoring processes improving more strongly across age than control processes (Krebs & Roebers, 2010, 2012). Similar patterns have been observed in science test

contexts (Roebers et al., 2009). Even younger children (5–7 years) demonstrate emerging metacognitive abilities: they can monitor the accuracy of their memory and, increasingly with age and ability, use these judgments to regulate behaviour such as study time or answer selection (Destan et al., 2014; Hembacher & Ghetti, 2014). Observational data further suggest that while monitoring improves primarily with age, control processes depend on both age and domain-specific competence (Bryce & Whitebread, 2012). Overall, metacognitive skills show a gradual developmental trajectory from preschool through late childhood, with early forms of introspective accuracy and strategic regulation already detectable but becoming more consistent and sophisticated over time. We can therefore assume that children have the basic metacognitive resources necessary to monitor their spelling by the end of primary school.

The present study

In this study, we investigated children's capacity to identify and self-correct spelling errors in their own writing as a window into metacognitive monitoring and control processes in early literacy acquisition. We were specifically interested in: (a) how accurate primary school children are in their metacognitive monitoring of their own writing of a dictated text, and (b) whether they are able to translate their spelling performance monitoring into regulatory behaviour such that they revise and correct detected spelling errors. While previous research provides evidence that some degree of monitoring and error-correction occurs in primary school children's writing, we seek to provide a clearer picture by differentiating between processes of metacognitive monitoring and control. The important aspect of our experimental design is that children not only monitored their spelling, but also attempted to correct their own spelling mistakes (i.e., without correct answer sheets, as in Cordewener et al., 2016 and Cordewener et al., 2018).

It is plausible that children encounter difficulties in self-correcting their own writing due to deficits in monitoring processes, control processes, and/or relevant domain-specific knowledge. Previous studies, for instance, suggest that children are able to accurately judge their performance in vocabulary and reading comprehension tasks, but that accurate performance monitoring does not necessarily then translate into regulatory behaviour in the form of study time allocation (e.g., Cornoldi, 1990; Metcalfe & Finn, 2013). Whether a similar deficit exists for spelling error detection and correction is unclear.

The literature on students' revisions of their initial answers on assessments also strongly suggests that school and university students are very cautious about making changes, preferring rather to stick to their initial answers (e.g., Benjamin et al., 1984; Higham & Gerrard, 2005; Tiffin-Richards et al., 2022; Waddell & Blankenship, 1994). This reluctance to change many initial answers, also referred to as a *first-instinct fallacy* (Kruger et al., 2005; Stone et al., 2022), is thought to be in part due to the desire to avoid the regret associated with making right-to-wrong changes (Couchman et al., 2016; Stylianou-Georgiou & Papanastasiou, 2017).

According to stage models of literacy development (Ehri, 2022; Frith, 1985), accurate spelling increasingly relies on orthographic strategies (Frith, 1985; consolidated alphabetic phase, Ehri, 2022), drawing on rule-based knowledge and representations stored in the mental lexicon (Perfetti, 2017). However, these models are best understood as describing grad-

ual and overlapping developmental processes rather than discrete stages. In early literacy, spelling performance is often relatively unstable, with children producing different spellings of the same word across occasions. While this variability likely reflects underspecified orthographic representations (see Endlich et al., 2025, for a discussion), it may also arise from the use of competing strategies (e.g., phonological or morphological encoding) as well as task-related demands. To our knowledge, no studies have yet addressed the issue of whether and how inconsistent spelling and revision behaviour interact.

During primary school, children learning alphabetic languages increasingly draw on orthographic strategies in writing, expand their sight vocabulary, and develop greater fluency and automaticity in reading (for children learning to read and spell in German, see Hasenäcker & Schroeder, 2022; Heß et al., 2026; Schneider, 2008) and as a result, spelling tends to become more stable over time. These developments may make it more likely that children detect errors when reviewing their own dictations. It is therefore plausible that years of formal literacy education, and the associated growth in language ability—particularly orthographic knowledge—contribute to children's metacognitive monitoring and control during self-correction.

To investigate children's ability to detect, revise, and correct spelling mistakes in their own writing, we conducted a standardized spelling assessment, as typically performed in primary schools in Germany: children wrote a continuous text dictated to them by an instructor and were then able to first review and then revise their writing. To be able to differentiate between the processes of monitoring and control, we explicitly instructed children to first underline words if they were unsure whether they had spelled them correctly, before making any changes. Only after the initial monitoring was completed (Review Phase) were children instructed that they could change their initial spellings (Revise Phase). This allowed us to separately analyse which spellings children were unsure of, representing their metacognitive monitoring of their own spelling, and their subsequent spelling revisions, representing their metacognitive regulation and revision processes.

Previous findings suggest that children tend to miss mistakes, rather than to falsely identify correctly spelled words as misspelled in error detection tasks (Endlich et al., 2021). In terms of SDT, we therefore expected values of d' significantly above 0, indicating differentiation between correctly and incorrectly spelled words (Hypothesis 1, H1). We expected *high diagnostic specificity* (>0.90), such that words that children rarely perceive correctly spelled words to contain spelling errors (H1.1), but relatively *low diagnostic sensitivity* (<0.5) to spelling errors, such that they miss the majority of their misspelled words (H1.2).

Furthermore, if children use their orthographic, morphological, and spelling rule knowledge to make judgements about spelling correctness and to guide self-corrections in their own writing, language proficiency and years of schooling should be positively related to monitoring processes. We therefore predicted that children's language-related school achievement and grade level, as well as their self-rated spelling ability would be positively associated with their ability to differentiate between correct and incorrect spellings (H2). We included both an objective measure (school grades) and subjective measure (self-rated spelling ability) of language performance, to capture a broad picture of children's language ability.

Finally, we know from the answer-changing literature that students stick to most of their initial answers, presumably to avoid making right-to-wrong changes (Higham & Gerrard, 2005). At the same time, answer changing is generally beneficial with more wrong-to-right

changes than right-to-wrong changes, resulting in a positive net gain (Tiffin-Richards et al., 2022). Accordingly, we also tested whether revision generally led to more corrections of spelling mistakes than changes of correct initial spellings, which would indicate successful regulation and correction processes (H3). In addition, we examined whether children's spelling awareness had an influence on the number of spellings children changed (H4.1) and their probability of correcting initial spelling errors (H4.2). To control for further individual differences, we included children's gender, test motivation, and self-rated spelling ability in all analyses. Gender was included as a control variable, as large-scale educational assessments in Germany find significantly higher orthographic performance in girls compared to boys in 4th grade (Stanat et al., 2022). Test motivation was included to control for invested effort (Baumert & Demmrich, 2001).

Transparency and data availability statement

We describe the sample and procedure and report all data exclusions in detail. All data, measures, and analysis codes are available at https://osf.io/9nc3d. Data were analysed using R, version 4.3.3 (R Core Team, 2024). The study design and hypotheses were not preregistered.

Methods

Participants

We recruited 318 children from eleven grade 3 classes ($N = 180$, 48% female) and eight grade 4 classes ($N = 138$, 48% female). All children were attending state-run primary schools in the German federal state of Bavaria. A majority of the participating children spoke only German at home (grade 3 = 69%, grade 4 = 71%), about a quarter spoke German amongst other languages (grade 3 = 25%, grade 4 = 26%), and a minority did not speak German at home (grade 3 = 6%, grade 4 = 3%). All children spoke German at a level allowing them to fully participate in school curriculum, in particular regular German classes. It is important to note that spelling tuition in German primary school education generally focusses on the use of phonological, morphological, and grammatical principles and official guidelines in the German state in which testing was conducted include the aim that children "demonstrate an awareness of spelling in their own writing by checking their spelling independently and correcting themselves." (Staatsinstitut für Schulqualität und Bildungsforschung, 2026).

This study was ethically approved by an Institutional Review Board (2025-11/97). Participants were fully informed about the study's content and procedures, participated voluntarily, could withdraw at any time, and data were collected anonymously. To ensure full anonymity, we refrained from collecting identifying data such as children's date of birth, exact age, or first language. Furthermore, the study posed no foreseeable risk to participants (see the ethical guidelines of the German Psychological Society, DGPs). Permission to conduct the study during school hours was granted by the responsible Ministry of Education.

Measures

Spelling assessment Spelling ability was measured using a standardized German spelling test for school grades 3 and 4 (Deutscher Rechtschreibtest für das dritte und vierte Schuljahr; DERET 3–4+; Stock & Schneider, 2008). The test provides grade-level-appropriate versions of texts that are dictated to children sentence by sentence. We employed test Form A, comprising 80 and 92 words for grades 3 and 4, respectively. Because the content of the text versions is essentially the same, with the longer version containing a few extra embedded or subordinate clauses, 73 words of the test versions are identical. We included only the 73 identical words across test versions in our analyses. The reliability of the spelling assessment was high for grades 3 ($\alpha = 0.93$) and 4 ($\alpha = 0.91$).

Academic achievement Children reported their most recent German grade as an indicator of their general language proficiency. Grades were reverse scored, so that 1 represented the lowest and 6 the highest achievement ($M = 4.50$, $SD = 0.83$).

Spelling self-rating Children reported their self-perceived spelling ability ("How good are you at spelling?"; German "Wie gut kannst du Rechtschreiben?") on a scale of 1 ("very poor"; German "gar nicht gut") to 5 ("very good"; German "sehr gut") with an age-appropriate single-item measure ($M = 3.43$, $SD = 0.86$).

Test motivation Children reported their state test motivation ("How important was it to you today to be as good as possible?"; German "Wie wichtig war es dir heute, möglichst gut zu sein?") on a scale of 1 ("not at all"; German "gar nicht wichtig") to 5 ("very important"; German "sehr wichtig") with an age-appropriate single-item measure ($M = 3.83$, $SD = 0.96$).

Procedure

Testing was conducted in classrooms at the participating schools. Trained instructors conducted the standardized procedure of the DERET 3–4+ (Stock & Schneider, 2008). After the dictation, the children were instructed to underline any word in their own written texts that they were not sure they had spelled correctly. In this *Review Phase*, no changes were made to the dictated text. Children used a green pen for the underlining stage. In the *Revise Phase*, children were then instructed that they could change any spelling in their own texts, using a green pen. Each phase lasted about 5 min. The testing sessions, including instructions (see Table 1) and short breaks, lasted 45 min. Children completed a brief demographic questionnaire, including self-rated spelling ability, test motivation, and their last German school grades.

Results

Accuracy on the 73 common words across test versions for grades 3 and 4 was high in grades 3 ($M = 70\%$, $SD = 15\%$, range = 27–99%) and 4 ($M = 82\%$, $SD = 12\%$, range = 44–100%). The results of the full DERET 3–4 + assessment indicated that 14% of the participating children

Table 1 Task instructions

Phase	Instruction
Introduction	"We'd like to do a few writing exercises with you today. But there won't be any grades. You won't even write your name on the worksheet. But do try your very best, because we want to see how well you can already write."
Write	standardized DERET 3–4 dictation instructions
Review	"Now please put your pen down and pick up a green pen. Do you have a green pen? Now please read through what you've written. Underline all the words you're not sure are correct. Please just underline them – don't correct anything yet."
Revise	"Now you can look through everything again, especially the underlined words, and think about whether you want to correct them."

The original instructions were given in German

in grade 3, and 15% in grade 4 had spelling ability falling below the 16th percentile. The extent of missing data was negligible (0–4.5% per variable, < 2% per child).

The number of spelling errors detected (Hits), errors missed (Misses), correct spellings falsely detected as misspelled (False Alarms), correct spellings that were not underlined (Correct Rejections) in the Review Phase are summarized in Table 2. All descriptive tables and analyses are based on the common 73 words across spelling assessments for grades 3 and 4. Children selected only a minority of their spellings for revision in the Review Phase (grade 3: 9%; grade 4: 5%). In the Review Phase, correctly spelled words that were not underlined (grade 3: 67%; grade 4: 80%) and incorrectly spelled words that were not underlined (grade 3: 24%; grade 4: 15%) occurred most frequently. Across grades, hits were twice as frequent as false alarms (grade 3: 6% vs. 3%; grade 4: 4% vs. 2%). With respect to the 73 words analysed, this translates into children, on average, correctly detecting 4 of their own spelling mistakes ($SD = 3$, range = 0–20), misidentifying 2 of their own correctly spelled words as errors ($SD = 2$, range = 0–13), missing 15 of their own spelling errors ($SD = 10$, range = 0–53), and correctly identifying 53 of their written words as correctly spelled ($SD = 11$, range = 19–73).

Metacognitive monitoring: spelling awareness

The number of correctly identified spelling mistakes (Hits) and correct spellings incorrectly identified as spelling mistakes (False Alarms) in the Review Phase were summed for each child and used to calculate their discrimination d' scores (see Cordewener et al., 2018, for a similar procedure). A log-linear correction was applied to each cell frequency in the contingency table by adding 0.5 to each cell, which corrects for extreme values of d' (Hautus, 1995). Individual d' values were estimated for each participant. Children's mean d' score, $M = 1.11$, $SD = 0.57$, was significantly above zero, $t(317) = 34.48$, $p < .001$, indicating that their discrimination of correctly and incorrectly spelled words was above chance level. This pattern of results supported our prediction that children can, in principle, discriminate between correct and incorrect spellings in their own writing (H1).

Children's sensitivity to their own spelling errors was low (H1.2), indicated by children in grade 3 only detecting 22% ($SD = 17\%$, range = 0–79%) of their spelling mistakes, and children in grade 4 similarly only detecting 23% ($SD = 10\%$, range = 0–100%) of their mistakes. Sensitivity was significantly below 0.5 for children in grades 3 and 4 (all p's < .001). In contrast, the specificity of children's error detection was very high (H1.1), indicated

Table 2 Frequencies, percentages, and correlations of correctly and incorrectly selected spelling in the review phase

Variable	*N*	%	1	2	3	*N*	%	1	2	3
	Grade 3					Grade 4				
All spellings	13,140					10,074				
Selected for revision (underlined)	1125	9				549	5			
1. Hits (H)	784	6				377	4			
2. False Alarms (FA)	341	3	.34***			172	2	.40***		
Not selected for revision										
3. Misses (Miss)	3192	24	.02	−.36***		1479	15	.22**	−.19*	
4. Correct Rejections (CR)	8823	67	−.41***	.03	−.90***	8046	80	−.60***	−.16	−.90***

Hits: selected words with spelling errors, False alarms: selected words without spelling errors, Misses: unselect words with spelling errors, Correct rejections: unselected words without spelling errors. Percentages are of the total number of spellings. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

by children in grade 3 judging 96% (*SD* = 5%, range = 72–100%) of their correctly spelled words to be correct, and children in grade 4 judging 98% (*SD* = 3%, range = 79–100%) of their correct spellings to be correct. Specificity was significantly above 0.9 for children in grades 3 and 4 (all *p*'s < .001).

The associations of participants' academic achievement, gender, self-rated spelling ability, and test motivation with their spelling awareness (d' scores) are summarized in the overview of their correlations in Table 3, and regression results in Table 4. Consistent with the idea that metacognitive spelling awareness is associated with general language skills, our analyses using d' scores as outcome variable indicated that children's ability to discriminate between their own correct and incorrect spellings increased as a function of their language proficiency (H2). This was supported by the significant positive association between children's d' scores and their German school grades, $b = 0.12$, $SE = 0.03$, $t = 3.68$, $p < .001$. Furthermore, metacognitive spelling awareness was significantly higher for children in grade 4 than for children in grade 3, $b = 0.14$, $SE = 0.04$, $t = 3.38$, $p < .001$, suggesting that awareness develops with age, language exposure, and schooling. Girls had significantly higher metacognitive spelling awareness than boys, $b = -0.07$, $SE = 0.03$, $t = -2.16$, $p = .032$.

We inspected the clustering effect of children nested within classes and schools, which revealed that only a small amount of variance in children's sensitivity, specificity, and spelling awareness was explained by class membership (5–7%). A robustness check using linear mixed effects models including random effects for class as grouping factor found no substantive changes to the pattern of results (see https://osf.io/9nc3d).

Table 3 Summary statistics from the signal detection theory analyses

Variable	1	2	3	4	5
1. Spelling awareness (*d'*)	—				
2. German grade	.28 ***	—			
3. Grade (3rd vs. 4th)	.20 ***	.02	—		
4. Self-rated spelling ability	.22 ***	.45***	.04	—	
5. Test motivation	.06	.01	−.11*	.09	—
6. Gender [a]	.17**	.11	< .01	.07	.19***

[a] gender was dummy coded 1 = girls, 0 = boys

*p < .05. **p < .01. ***p < .001

Table 4 Predicting children's spelling awareness (*d'*)

Intercept	b	SE	t-value	*p*-value
	1.15	0.03	36.98	< .001
Grade (3rd vs. 4th)	0.14	0.04	**3.38**	< .001
German grade	0.12	0.03	**3.68**	< .001
Self-rated spelling ability	0.08	0.04	**1.97**	.050
Test motivation	0.03	0.03	0.88	.380
Gender [a]	−0.07	0.03	**−2.16**	.032

[a] gender was effect coded 1 = girls, −1 = boys

Metacognitive control: revision and self-correction

The number of revised spellings and wrong-to-right, wrong-to-wrong, and right-to-wrong changes are summarized for children in grades 3 and 4 in Table 5. Children were clearly able to detect which words they were likely to have misspelled by underlining them, with 71% of the underlined words being misspelled. However, they did not attempt to correct all of the words they identified as problematic, changing only 24% of the underlined words. Of the underlined words that were subsequently revised, the majority (averaged across both grades) were wrong-to-wrong revisions (48%), over a third were successful wrong-to-right corrections (41%), and only a minority were right-to-wrong revisions (11%). With respect to the 73 words analysed, this translates into children, on average, correcting 0.8 spelling mistakes ($SD = 1.1$, range = 0–7), unsuccessfully changing the spelling of 0.9 incorrectly spelled words ($SD = 1.4$, range = 0–7), and changing the spelling of 0.2 words with correct spelling, resulting in an incorrect spelling ($SD = 0.54$, range = 0–4). A net gain of revision was calculated for each child by scoring wrong-to-right revisions as +1, wrong-to-wrong revisions as 0, and right-to-wrong revisions as −1. Revising spellings thus lead to a net gain of 0.6 correct spellings per child. A regression model with net gain as a continuous dependent variable resulted in a significant positive effect of revision, $b = 0.23$, $SE = 0.03$, $t = 7.06$, $p < .001$, which supports Hypothesis 3.

To test whether children's spelling awareness had an influence on the number of spellings they changed (H4.1) and their probability of correcting initial spelling errors (H4.2), we ran two additional regression models. In the first, a generalized linear regression model was used with revision status (revised, not revised) as the binomial outcome variable. In the second model, we used a linear regression model with the *net gain* of revision as outcome variable. In both models, we included children's spelling discrimination (d'), academic achievement (continuous, centred), self-rated spelling ability (continuous, centred), test motivation (continuous, centred), gender (effect coded, girls = 1, boys = −1) and grade level (effect coded, grade 4 = 1, grade 3 = −1) as predictors. It is important to note that we only considered changes to words that were selected (i.e., underlined) in the Review Phase, so that children's spelling discrimination (d') could be used as a relevant predictor.

The results of the regression analyses are summarised in Table 6. Children with high spelling discrimination had higher odds of revising the spelling of a word that they had previously selected (i.e., underlined) in the Review Phase, $b = 0.61$, $SE = 0.13$, $z = 4.81$, $p < .001$. High test motivation was also associated with higher odds of revising a word selected in the Review Phase, $b = 0.25$, $SE = 0.07$, $z = 3.40$, $p < .001$. Girls also had higher odds of revising a previously selected word, $b = 0.16$, $SE = 0.07$, $z = 2.50$, $p = .013$. Concerning the net gain of revision, only children's spelling discrimination was positively associated with the success of revising a previously selected word, $b = 0.19$, $SE = 0.08$, $t = 2.50$, $p = .013$.

Table 5 Number and success of spelling revisions

Participants			Changes				
Grade	N_{part}	N_{resp}	$N_{changes}$	% changes	RW	WW	WR
3	180	13,140	384	3	11%	51%	38%
4	138	10,074	222	2	12%	43%	45%

N_{part} participants, N_{resp} responses (spelled words), *WR* wrong-to-right change, *WW* wrong-to-wrong change, *RW* right-to-wrong change

Table 6 Analyses of revisions and revision success

	b	SE	t-value	*p*-value
	Revision of selected words			
Intercept	−1.99	0.18	**−10.94**	<.001
d'	0.63	0.13	**4.81**	<.001
Grade (3rd vs. 4th)	0.01	0.07	0.16	.876
German grade	−0.07	0.09	−0.78	.435
Gender	0.16	0.07	**2.50**	.013
Self-rated spelling ability	−0.08	0.08	−0.99	.323
Test motivation	0.25	0.07	**3.40**	<.001
	Revision success			
Intercept	−0.02	0.11	−0.16	.871
d'	0.19	0.08	**2.50**	.013
Grade (3rd vs. 4th)	−0.05	0.04	−1.33	.183
German grade	0.08	0.05	1.54	.124
Gender [a]	0.06	0.04	1.57	.118
Self-rated spelling ability	0.03	0.05	0.53	.595
Test motivation	−0.07	0.04	−1.70	.091

[a] gender was effect coded 1 = girls, −1 = boys

Discussion

Our study aimed to use children's capacity to identify and self-correct spelling errors in their own writing as a window into metacognitive monitoring and control processes in early literacy acquisition. Our results demonstrate that children attending German primary school are to some extent capable of detecting errors in their own writing. They do so with a low detection rate, meaning that they miss the majority of their spelling errors (low sensitivity), but rarely mistakenly perceive an error in correctly spelled words (high specificity), leading to a high discrimination rate between their correct and incorrect spellings. Our results also indicate that when children revise their initial spellings, they (a) do so very hesitantly, only attempting to revise 24% of the words they identify as possibly containing a spelling error, and (b) do so with higher odds of correcting an initial spelling mistake than changing a correct spelling.

Our synthesis of the metacognitive monitoring and answer-changing literature suggested three challenges for successful error revision: (1) accurate performance monitoring (metacognitive monitoring); (2) translation of monitoring judgements into regulatory behaviour, such as answer revision (metacognitive control); and (3) the domain-specific or procedural knowledge necessary to retrieve the correct answer from memory or solve the problem at hand. According to the *implementation deficit* account (Metcalfe, 2009), metacognitive processes often fail because accurate performance monitoring does not translate into efficient behavioural strategies and regulation (Cornoldi, 1990; Dufresne & Kobasigawa, 1989; Koriat et al., 2009). However, our results suggest that children may already face difficulties at the monitoring stage and that these difficulties are related to their domain-specific knowledge.

In our study, children missed most of the mistakes that they made while writing the text dictated to them by an instructor. It is possible that children did not underline some words they thought were incorrectly spelled because they did not know how to correct them and therefore saw no need or point to mark them. However, the task instructions in the Review Phase were intended to avoid this, as children were initially simply asked to underline words

with spelling patterns they were unsure of. Only in the subsequent Review Phase were children instructed that they could change the spelling of any of the words in their written dictation. This suggests that children failed to underline many of their spelling mistakes because they lacked the awareness (i.e., monitoring capacity) to recognize them as such.

Our findings also suggest that metacognitive monitoring of spelling is influenced by domain-specific knowledge. Children's ability to judge the correctness of their own spellings was positively associated with three factors: their academic grades in German (their language of instruction), their grade level (3rd vs. 4th), and their self-reported spelling ability. These associations indicate that metacognitive monitoring is related to both language proficiency and years of formal literacy education. This supports the view that monitoring processes are grounded in domain-specific knowledge—in this case, the quality of orthographic representations in the mental lexicon (Perfetti, 2017) and knowledge of spelling rules—which may facilitate accurate spelling awareness during writing.

Children's difficulties in using self-correction strategies in their own writing therefore appear to be related, at least in part, to limitations of their metacognitive monitoring of their own spelling. This is not to say that these children do not also experience an implementation deficit (Metcalfe, 2009). Indeed, a clear result of our study is children's reluctance to attempt to revise spellings they assume to contain a spelling error. However, children missed most of their mistakes before even attempting to revise any of their identified mistakes.

Practical implications for self-correction

Against the backdrop of integrated spell-checking and auto-correction in digital devices, spelling might seem largely irrelevant as an educational goal. However, it would be wrong to surmise that spelling ability has lost its relevance in todays' society or as an important indicator of language proficiency (Pan et al., 2021). Spelling difficulties can bind processing capacities during writing (Graham & Santangelo, 2014) and lead to children viewing themselves as poor writers, instead of poor spellers (Graham, 1999), with detrimental effects on their further writing development. Improving children's ability to self-correct their writing may help counteract these negative impacts on writing skills and development.

The central finding of our study is that children's self-correction of dictated texts is constrained both by limitations in spelling monitoring and by hesitancy or difficulty in correcting misspelled words. However, when children do engage in revision, they are generally successful in correcting errors. Given that spelling instruction typically already incorporates phonological, morphological, and grammatical principles (as in the classrooms in the present study), it is plausible that monitoring, control, and revision processes could be further supported through more explicit and systematic use of these established instructional practices. For example, applying syllable structure approaches (e.g., Müller et al., 2020; see also Cordewener et al., 2016, 2018) or spelling rules based on phoneme–grapheme correspondences (i.e., phonics instruction) and morphology (Treiman, 2018) may enhance children's sensitivity to implausible spelling patterns and typographical errors, while reinforcing spelling conventions (although no long-term benefits were found in Cordewener et al., 2018, using correct answer sheets). These approaches align with established interventions combining phonics, orthographic, and morphological instruction, which have proven effective for learners with developmental dyslexia (see Galuschka et al., 2020, for a meta-analysis).

A complementary approach would be to explicitly teach children the value of checking and revising their work, which may help reduce hesitancy to correct initial responses. Even minimal interventions of this kind have been shown to produce small increases in revision behaviour in university students (Tiffin-Richards et al., 2022). Combining such approaches with adaptive, digital, game-based spelling programs may offer additional benefits, as these can be tailored to individual performance levels and used to supplement standard spelling instruction (see Witzel et al., 2024).

Limitations and future directions

Our study design allowed us to distinguish children's monitoring processes from control and revision processes by requiring them to underline words they were unsure of before attempting to revise them. We do not, however, know whether children considered revising a word but then decided against making a change. We also cannot say whether revisions were not attempted because of a failure to translate successful detection of an error (i.e., successful metacognitive monitoring) or whether children lacked the necessary orthographic knowledge to produce the correct spelling. Failure to change a detected error or failure to successfully revise an error may therefore either be attributed to an implementation deficit (Metcalfe, 2009) or lack of orthographic knowledge (Landerl et al., 2022). It is therefore not possible to definitively distinguish between implementation deficits and (strategic) knowledge deficits.

Furthermore, we did not collect a measure of certainty for misspelled words that were not selected for revision (i.e., underlined in the Review Phase). We therefore do not know for certain whether when children did not underline a misspelled word (signifying a miss in SDT) because they had an inaccurate mental orthographic representation of the word, or simply did not detect the error. To be able to distinguish between these possibilities would require experimental designs such as individual ratings of spelling certainty for all written words (i.e., on a scale rather than underlying "uncertain" words, as in our study), or qualitative explanations why specific words were or were not revised. Low confidence misses could indicate low quality mental representations of a word and "guessing", whereas high confidence misses could indicate inaccurate mental representations or inappropriately applied spelling rules.

A related limitation is that we did not directly assess children's vocabulary or orthographic knowledge and relied on more general language ability measures including German language school grades and self-reported spelling ability, as well as the years of formal education. However, the fact that we found significant associations between these distal measures of children's spelling ability and their ability to discriminate between correct and incorrect spelling in their own writing suggest that more direct measures of orthographic knowledge may find even larger effects of domain-specific knowledge on metacognitive monitoring.

A key methodological distinction between our study and traditional error detection tasks (see Endlich et al., 2025) is that, rather than evaluating predetermined examples of correct and incorrect spelling patterns, participants in our study produced both correct and incorrect spellings themselves. Consequently, the type and frequency of spelling errors varied among participants. However, to ensure comparability across individuals, children wrote age-appropriate versions of the same dictated text that differed only in a small number of words. This design may more closely reflect authentic spelling processes in school contexts and provide avenues for formative assessment, using children's own spelling mistakes to generate subsequent learning materials or prompt relevant spelling rules. An open question for future research is

whether metacognitive monitoring operates differently in free writing tasks, where children generate their own lexical choices and may rely on different strategies for error detection—in particular the avoidance of writing words they are not sure they can spell correctly.

Conclusion

Based on our findings, we draw several conclusions. First, primary school children often fail to detect their own spelling errors, suggesting a lack of spelling awareness rather than a lack of correction skill per se. Second, even when they suspect a word may be misspelled, children are frequently reluctant to revise it. Third, when revisions are attempted, they are often successful, indicating that the ability to correct exists but is underutilized. This pattern suggests that self-correction during dictation tasks is unlikely to be effective without explicit instructional scaffolding and support. Targeted teaching of metacognitive strategies may help children to better monitor and revise uncertain spellings, especially in cases where they cannot retrieve the exact orthographic from memory. Further training studies are needed to determine how increases in revision behaviour translate into greater effectiveness, for example by improving the quantity and quality of corrections made.

Author contributions **SimonTiffin-Richards:** Conceptualization; Data curation; Formal analysis; Writing - original draft; Writing - review & editing. **Jan Lenhart**: Conceptualization; Validation; Writing - review & editing. **Peter Marx**: Conceptualization; Resources; Data curation; Investigation; Writing - review & editing.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Competing interest The authors report there are no competing interests to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Baker, L. (1989). Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review, 1*(1), 3–38. https://doi.org/10.1007/BF01326548

Baumert, J., & Demmrich, A. (2001). Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education*, 16, 441–462. https://doi.org/10.1007/BF03173192

Benjamin, A. S., & Diaz, M. (2008). Measurement of relative metamnemonic accuracy. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of memory and metamemory* (pp. 73–94). Psychology Press.

Benjamin, L. T., Cavell, A., & Shallenberger, W. R. (1984). Staying with initial answers on objective tests: Is it a myth? *Teaching of Psychology, 11*(3), 133–141. https://doi.org/10.1177/0098628384011003

Block, K. K., & Peskowitz, N. B. (1990). Metacognition in spelling: Using writing and reading to self-check spellings. *The Elementary School Journal, 91*(2), 151–164. https://doi.org/10.1086/461643

Boekaerts, M., & Rozendaal, J. S. (2010). Using multiple calibration indices in order to capture the complex picture of what affects students' accuracy of feeling of confidence. *Learning and Instruction, 20*, 372–382. https://doi.org/10.1016/j.learninstruc.2009.03.002

Bryce, D., & Whitebread, D. (2012). The development of metacognitive skills: Evidence from observational analysis of young children's behavior during problem-solving. *Metacognition and Learning, 7*, 197–217. https://doi.org/10.1007/s11409-012-9091-2

Cordewener, K. A., Verhoeven, L., & Bosman, A. M. (2016). Improving spelling performance and spelling consciousness. *The Journal of Experimental Education, 84*(1), 48–74. https://doi.org/10.1080/00220973.2014.963213

Cordewener, K. A., Hasselman, F., Verhoeven, L., & Bosman, A. M. (2018). The role of instruction for spelling performance and spelling consciousness. *The Journal of Experimental Education, 86*(2), 135–153. https://doi.org/10.1080/00220973.2017.1315711

R Core Team (2024). *R: A language and environment for statistical computing (Version 4.0.0)* [Computer software]. R Foundation for Statistical Computing. http://www.R-project.org

Cornoldi, C. (1990). Metacognitive control processes and memory deficits in poor comprehenders. *Learning Disability Quarterly, 13*(4), 245–255. https://doi.org/10.2307/151035

Couchman, J. J., Miller, N. E., Zmuda, S. J., Feather, K., & Schwartzmeyer, T. (2016). The instinct fallacy: The metacognition of answering and revising during college exams. *Metacognition & Learning, 11*(2), 171–185. https://doi.org/10.1007/s11409-015-9140-8

De Boer, H., Donker, A. S., Kostons, D. D., & Van der Werf, G. P. (2018). Long-term effects of metacognitive strategy instruction on student academic performance: A meta-analysis. *Educational Research Review, 24*, 98–115. https://doi.org/10.1016/j.edurev.2018.03.002

Destan, N., Hembacher, E., Ghetti, S., & Roebers, C. M. (2014). Early metacognitive abilities: The interplay of monitoring and control processes in 5-to 7-year-old children. *Journal of Experimental Child Psychology, 126*, 213–228. https://doi.org/10.1016/j.jecp.2014.04.001

Donker, A. S., De Boer, H., Kostons, D., Van Ewijk, C. D., & van der Werf, M. P. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review, 11*, 1–26. https://doi.org/10.1016/j.edurev.2013.11.002

Dufresne, A., & Kobasigawa, A. (1989). Children's spontaneous allocation of study time: Differential and sufficient aspects. *Journal of Experimental Child Psychology, 47*(2), 274–296. https://doi.org/10.1016/0022-0965(89)90033-7

Ehri, L. C. (2022). What teachers need to know and do to teach letter–sounds, phonemic awareness, word reading, and phonics. *The Reading Teacher, 76*(1), 53–61. https://doi.org/10.1002/trtr.2095

Endlich, D., Richter, T., Marx, P., Lenhard, W., Moll, K., Witzel, B., & Schulte-Körne, G. (2021). Spelling error detection: A valid and economical task for assessing spelling skills in elementary-school children. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 52(*1–2), 25–40. https://doi.org/10.1026/0049-8637/a000227

Endlich, D., Lenhard, W., Marx, P., & Richter, T. (2025). Measuring Spelling Skills: A Meta-Analysis of the Convergence of Dictation and Error Detection Tasks. *Educational Research Review*, Article 100706. https://doi.org/10.1016/j.edurev.2025.100706

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience, 8*, Article 443. https://doi.org/10.3389/fnhum.2014.00443

Fleming, S. M. (2024). Metacognition and confidence: A review and synthesis. *Annual Review of Psychology, 75*(1), 241–268. https://doi.org/10.1146/annurev-psych-022423-032425

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia, Neuropsychological and Cognitive Studies of Phonological Reading* (pp. 301–330). Erlbaum.

Galuschka, K., Görgen, R., Kalmar, J., Haberstroh, S., Schmalz, X., & Schulte-Körne, G. (2020). Effectiveness of spelling interventions for learners with dyslexia: A meta-analysis and systematic review. *Educational Psychologist, 55*(1), 1–20. https://doi.org/10.1080/00461520.2019.1659794

Georghiades, P. (2004). From the general to the situated: Three decades of metacognition. *International Journal of Science Education, 26*(3), 365–383. https://doi.org/10.1080/0950069032000119401

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing, 27*, 1703–1743. https://doi.org/10.1007/s11145-014-9517-0

Graham, S. (1999). Handwriting and spelling instruction for students with learning disabilities: A review. *Learning Disability Quarterly, 22*, 78–98. https://doi.org/10.2307/1511268

Green DM, Swets JA (1966) *Signal detection theory and psychophysics*. John Wiley
Grothe, B. F. (1966). A Study of Spelling Conscience. *Elementary English, 43*(7), 774–784. https://www.jstor.org/stable/41386075
Hasenäcker, J., Schroeder, S. (2022). Specific predictors of length and frequency effects in German beginning readers: testing component processes of sublexical and lexical reading in the DRC. *Reading and Writing 35*, 1627–1650. https://doi.org/10.1007/s11145-021-10251-5
Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behavior Research Methods, Instruments, & Computers, 27*(1), 46–51. https://doi.org/10.3758/BF03203619
Hembacher, E., & Ghetti, S. (2014). Don't look at my answer: Subjective uncertainty underlies preschoolers' exclusion of their least accurate memories. *Psychological Science, 25*(9), 1768–1776. https://doi.org/10.1177/0956797614542273
Heß, J., Karageorgos, P., Endlich, D., & Richter, T. (2026). From sublexical to lexical processing: Developmental effects of word characteristics on word recognition difficulty in a transparent orthography. *Scientific Studies of Reading*, *30*(1), 1–28. https://doi.org/10.1080/10888438.2025.2586600
Higham, P. A., & Gerrard, C. (2005). Not all errors are created equal: Metacognition and changing answers on multiple-choice tests. *Canadian Journal of Experimental Psychology, 59*(1), 28–34. https://doi.org/10.1037/h0087457
Kadlec, H. (1999). Statistical properties of d' and estimates of signal detection theory. *Psychological Methods, 4*(1), 23e43. https://doi.org/10.1037/1082-989X.4.1.22
Koriat, A. (2012). When are two heads better than one and why? *Science, 336*(6079), 360–362. https://doi.org/10.1126/science.1216549
Koriat, A., Ackerman, R., Lockl, K., & Schneider, W. (2009). The memorizing effort heuristic in judgments of learning: A developmental perspective. *Journal of Experimental Child Psychology, 102*(3), 265–279. https://doi.org/10.1016/j.jecp.2008.10.005
Koriat, A. (2008). Subjective confidence in one's answers: The consensuality principle. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 945–959. https://doi.org/10.1037/0278-7393.34.4.945
Krebs, S. S., & Roebers, C. M. (2010). Children's strategic regulation, monitoring, and control skills during test-taking: The role of retrieval processes and scoring scheme. *British Journal of Educational Psychology, 80*, 325–340. https://doi.org/10.1348/000709910X485719
Krebs, S. S., & Roebers, C. M. (2012). The impact of retrieval processes, age, general achievement level, and test scoring scheme for children's metacognitive monitoring and controlling. *Metacognition and Learning, 7*(2), 75–90. https://doi.org/10.1007/s11409-011-9079-3.
Kreiner, D. S., & Green, A. G. (2000). Feeling-of-knowing judgments and strategy selection in spelling. *Perceptual and Motor Skills, 90*, 775–786. https://doi.org/10.2466/pms.2000.90.3.775
Kruger, J., Wirtz, D., & Miller, D. T. (2005). Counterfactual thinking and the first instinct fallacy. *Journal of Personality and Social Psychology, 88*(5), 725–735. https://doi.org/10.1037/0022-3514.88.5.725
Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science, 9*(5), 178–181. https://doi.org/10.1111/1467-8721.0008
Landerl, K., Castles, A., & Parrila, R. (2022). Cognitive precursors of reading: A cross-linguistic perspective. *Scientific Studies of Reading, 26*(2), 111–124. https://doi.org/10.1080/10888438.2021.1983820
Lingel, K., Lenhart, J., & Schneider, W. (2019). Metacognition in mathematics: Do different metacognitive monitoring measures make a difference? *ZDM Mathematics Education, 51*(4), 587–600. https://doi.org/10.1007/s11858-019-01062-8
Lull, H. G. (1917). A plan for developing a spelling consciousness. *The Elementary School Journal, 17*(5), 355–361.
Lynn, S. K., & Barrett, L. F. (2014). "Utilizing" signal detection theory. *Psychological Science, 25*(9), 1663–1673. https://doi.org/10.1177/095679761454199
McFarland, W. H. (1916). Relation between spelling judgment and spelling ability. *Midland Schools, 31*, 19–21.
McGuffin, M. E., Martz, S. A., & Heron, T. E. (1997). The effects of self-correction versus traditional spelling on the spelling performance and maintenance of third grade students. *Journal of Behavioral Education, 7*, 463–476.
Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science, 18*(3), 159–163. https://doi.org/10.1111/j.1467-8721.2009.01628.x
Metcalfe, J., & Finn, B. (2013). Metacognition and control of study choice in children. *Metacognition and Learning, 8*(1), 19–46. https://doi.org/10.1007/s11409-013-9094-7
Müller, B., Richter, T., & Karageorgos, P. (2020). Syllable-based reading improvement: Effects on word reading and reading comprehension in grade 2. *Learning and Instruction, 66*, Article 101304. https://doi.org/10.1016/j.learninstruc.2020.101304

Nelson, T. O. (1990). Metamemory: A theoretical framework and new findings. *Psychology of learning and motivation* (Vol. 26, pp. 125–173). Academic Press.
Nietfeld, J. L., Cao, L., & Osborne, J. W. (2005). Metacognitive monitoring accuracy and student performance in the postsecondary classroom. *The Journal of Experimental Educational, 74*(1), 7–28.
Pan, S. C., Rickard, T. C., & Bjork, R. A. (2021). Does spelling still matter—and if so, how should it be taught? Perspectives from contemporary and historical research. *Educational Psychology Review, 33*(4), 1523–1552. https://doi.org/10.1007/s10648-021-09611-y
Pennequin, V., Questel, F., Delaville, E., Delugre, M., & Maintenant, C. (2020). Metacognition and emotional regulation in children from 8 to 12 years old. *British Journal of Educational Psychology, 90*, 1–16. https://doi.org/10.1111/bjep.12305
Perfetti, C. A. (2017). Lexical quality revisited. In E. Segers & P. van den Broek (Eds.), *developmental perspectives in written language and literacy: In honor of Ludo Verhoeven* (pp. 51–67). John Benjamins Publishing Company. https://doi.org/10.1075/z.206.04per
Pryor, H. C., & Pittman, M. S. (1921). *Guide to teaching spelling*. Macmillan Company.
Roebers, C. M., Schmid, C., & Roderer, T. (2009). Metacognitive monitoring and control processes involved in primary school children's test performance. *British Journal of Educational Psychology, 79*, 749–767. https://doi.org/10.1348/978185409X429842
Schneider, W. (2008). Entwicklung und Erfassung der Rechtschreibkompetenz im Jugend- und Erwachsenenalter [Development and assessment of spelling competence in youth and adulthood]. In W. Schneider, H. Marx, & M. Hasselhorn (Eds.), *Diagnostik von Rechtschreibleistungen und -kompetenz* (pp. 145–157). Hogrefe.
Schneider, W., & Artelt, C. (2010). Metacognition and mathematics education. *ZDM, 42*(2), 149–161.
Schraw, G., Kuch, F., & Gutierrez, A. P. (2013). Measure for measure: Calibrating ten commonly used calibration scores. *Learning and Instruction, 24*, 48–57. https://doi.org/10.1016/j.learninstruc.2012.08.007
Schraw, G., Kuch, F., Gutierrez, A. P., & Richmond, A. S. (2014). Exploring a three-level model of calibration accuracy. *Journal of Educational Psychology, 106*(4), 1192–1202. https://doi.org/10.1037/a0036653
Schraw, G. (2009). Measuring metacognitive judgments. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 415–429). Erlbaum.
Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033
Shaughnessy, J. J. (1979). Confidence-judgment accuracy as a predictor of test performance. *Journal of Research in Personality, 13*(4), 505–514.
Silva, C., Peixoto, F., & Salvador, L. (2021). To give someone a fish or teach them how to fish? Effects of a self-reflection tool on orthographic performance in Portuguese children. *Reading and Writing, 34*, 2481–2503. https://doi.org/10.1007/s11145-021-10154-5
Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(1), 204–221. https://doi.org/10.1037/0278-7393.26.1.204
Staatsinstitut für Schulqualität und Bildungsforschung (2026) *LehrplanPLUS*. Retrieved April 1, 2026, from https://www.lehrplanplus.bayern.de/fachlehrplan/grundschule
Stanat, P., Schipolowski, S., Schneider, R., Sachse, K. A., Weirich, S., & Henschel, S. (Eds.). (2022). *IQB-Bildungstrend 2021: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich*. Waxmann Verlag.
Stock, C., & Schneider, W. (2008). *Deutscher Rechtschreibtest für das dritte und vierte Schuljahr*. Hogrefe.
Stone, C., Mattingley, J. B., & Rangelov, D. (2022). On second thoughts: Changes of mind in decision-making. *Trends in Cognitive Sciences, 26*(5), 419–431. https://doi.org/10.1016/j.tics.2022.02.004
Stylianou-Georgiou, A., & Papanastasiou, E. C. (2017). Answer changing in testing situations: The role of metacognition in deciding which answers to review. *Educational Research and Evaluation, 23*(3–4), 102–118. https://doi.org/10.1080/13803611.2017.1390479
Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science, 240*(4857), 1285–1293. https://doi.org/10.1126/science.3287615
Tidyman, W. F. (1924). Do elementary school pupils know when they make mistakes in spelling. *School and Society, 20*, 349–350.
Tiffin-Richards, S. P., Lenhart, J., & Marx, P. (2022). When do examinees change their initial answers? The influence of task instruction, response confidence, and subjective task difficulty. *Learning and Instruction, 82*, 101654. https://doi.org/10.1016/j.learninstruc.2022.101654
Treiman, R. (2018). Teaching and learning spelling. *Child Development Perspectives, 12*(4), 235–239. https://doi.org/10.1111/cdep.12292

Tulis, M., Steuer, G., & Dresel, M. (2016). Learning from errors: A model of individual processes. *Frontline Learning Research, 4*(2), 12–26. https://doi.org/10.14786/flr.v4i2.168

Valmont, W. J. (1972). Spelling consciousness: A long neglected area. *Elementary English, 49*(8), 1219–1265.

Waddell, D. L., & Blankenship, J. C. (1994). Answer changing: A meta-analysis of the prevalence and patterns. *The Journal of Continuing Education in Nursing, 25*(4), 155–158.

Witzel, B., Görgen-Rein, R., Galuschka, K., Huemer, S., Del Toro, I. C., Schulte-Körne, G., & Moll, K. (2024). Digital game-based spelling intervention for children with spelling deficits: A randomized controlled trial. *Learning and Instruction*, *89*, 101842. https://doi.org/10.1016/j.learninstruc.2023.101842

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1594), 1310–1321. https://doi.org/10.1098/rstb.2011.0416

Zhang, Q., & Fiorella, L. (2023). An integrated model of learning from errors. *Educational Psychologist, 58*(1), 18–34. https://doi.org/10.1080/00461520.2022.2149525

Zohar, A., & Barzilai, S. (2013). A review of research on metacognition in science education: Current and future directions. *Studies in Science Education, 49*(2), 121–169. https://doi.org/10.1080/03057267.2013.847261

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.