COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 47 (2023) e13310

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13310

Limited Evidence of an Association Between Language, Literacy, and Procedural Learning in Typical and Atypical Development: A Meta-Analysis

Cátia M. Oliveira, Lisa M. Henderson, Marianna E. Hayiou-Thomas

Department of Psychology, University of York

Received 1 August 2022; received in revised form 6 June 2023; accepted 12 June 2023

Abstract

The ability to extract patterns from sensory input across time and space is thought to underlie the development and acquisition of language and literacy skills, particularly the subdomains marked by the learning of probabilistic knowledge. Thus, impairments in procedural learning are hypothesized to underlie neurodevelopmental disorders, such as dyslexia and developmental language disorder. In the present meta-analysis, comprising 2396 participants from 39 independent studies, the continuous relationship between language, literacy, and procedural learning on the Serial Reaction Time task (SRTT) was assessed across children and adults with typical development (TD), dyslexia, and Developmental Language Disorder (DLD). Despite a significant, but very small, relationship between procedural learning and overall language and literacy measures, this pattern was not observed at the group-level when examining TD, dyslexic, and DLD groups separately. Based on the procedural/declarative model, a positive relationship was expected between procedural learning and language and literacy measures for the typically developing group; however, no such relationship was observed. This was also the case for the disordered groups ($ps > .05$). Also counter to expectations, the magnitude of the relationship between procedural learning and grammar and phonology did not differ between TD and DLD ($ps > .05$), nor between the TD and dyslexic group on reading, spelling, and phonology ($ps >$ .05). While lending little support to the procedural/declarative model, we consider that these results

Marianna E. Hayiou-Thomas and Lisa M. Henderson share co-senior authorship.

Correspondence should be sent to Cátia M. Oliveira, Department of Psychology, University of York, Heslington, York, Y010 5DD, UK. E-mail: catia.oliveira@york.ac.uk

may be the consequence of poor psychometric properties of the SRTT as a measure of procedural learning.

Keywords: Procedural memory; Sequence learning; Individual differences; Language; Literacy; Dyslexia; Developmental language disorder; Specific language impairment

1. Introduction

Procedural learning refers to the ability to learn, consolidate, and control motor and cognitive skills that require the integration of statistical, probabilistic, and sequence knowledge (Batterink, Paller, & Reber, 2019; Packard & Knowlton, 2002; Ullman, Earle, Walenski, & Janacsek, 2020). This memory system has been proposed to support the development and acquisition of language (Ullman, 2004; Ullman et al., 2020), specifically for linguistic subdomains that require the extraction of patterns, such as phonology and grammar (Christiansen, Conway, & Onnis, 2012; Conway & Pisoni, 2008; Ullman, 2004; Ullman & Pierpont, 2005; Ullman et al., 2020). Deficits of procedural memory are also claimed to be a core causal factor in dyslexia and developmental language disorder (Ullman, 2004). In line with this account, there is clear evidence for a procedural deficit in populations with these neurodevelopmental disorders when compared to controls (e.g., Lum, Ullman, & Conti-Ramsden, 2013, 2014; West, Melby-Lervåg, & Hulme, 2021). In contrast, recent meta-analyses examining the continuous relationship between procedural learning and language and literacy found negligible and nonsignificant evidence for such an association (Lammertink, Boersma, Wijnen, & Rispens, 2020; West et al., 2021; cf. Hamrick, Lum, & Ullman, 2018 who did report a significant overall association, albeit on the basis of a much smaller sample of studies). However, the extant correlational meta-analyses have been limited as they mainly focused on grammar and vocabulary (Hamrick et al., 2018; Lammertink et al., 2020), with the exception of West et al. (2021) which combined measures of language and literacy. Therefore, here, we extend existing meta-analyses by producing the largest and most comprehensive meta-analysis examining the continuous relationship between procedural learning and different language and literacy subdomains, and directly comparing the magnitude of the relationship between language and literacy and procedural learning in typical and atypical populations with language-related learning difficulties (dyslexia and developmental language disorder).

1.1. Procedural learning in typical and atypical language and literacy development

Recent perspectives on language learning, including work in statistical learning, emphasize the importance of the detection and acquisition of regularities in the input as a fundamental aspect of language and literacy development (Bogaerts, Siegelman, & Frost, 2021; Saffran, 2020; Smith, Suanda, & Yu, 2014). One influential model, particularly in the area of developmental language and literacy disorders, is the Declarative/Procedural model developed by Ullman and colleagues, which highlights the role of the domain-general procedural learning system in the learning of these regularities (Ullman, 2001a, 2001b, 2004, 2016a, 2016b; Ullman et al., 2020; Ullman & Pierpont, 2005; Ullman & Pullman, 2015). According to this model, the procedural memory system is hypothesized to be involved in the acquisition and processing of syntax (the set of rules for how linguistic units are combined to convey meaning), morphology (how words and morphemes are combined to form words), and phonology (the abstract representations and rules that allow for the combination of sounds into more complex linguistic structures such as morphemes; van der Lely & Pinker, 2014). These fundamental domains of language are proposed to be acquired through gradual exposure to statistical, probabilistic, and sequential structures (Ullman, 2001a, 2001b, 2004, 2016a, 2016b; Ullman et al., 2020). Other aspects of language may also be learned through the procedural memory system, such as word boundaries in a speech stream, articulatory knowledge, and speech perception (Ullman et al., 2020). Conversely, the declarative memory system, which underpins the acquisition of arbitrary and idiosyncratic knowledge, is thought to be relevant to the accumulation of vocabulary knowledge fundamental to the mental lexicon.

However, given the learning flexibility of the declarative memory system, this long-term memory system is also expected to be involved in the acquisition of grammar through the learning and storage of the rules using associative learning mechanisms, such as chunking (Ullman & Pullman, 2015). Taking this into account, the Procedural/Declarative model currently lacks a precise delineation of the involvement of the procedural memory system in language and literacy development, as it is not clear whether procedural memory is expected to be involved in all aspects of syntactic, morphological, and phonological acquisition that involve pattern extraction. Furthermore, it may well be that rather than a direct mapping between one learning system and one behaviorally defined component of language (e.g., procedural learning and syntax; declarative learning and vocabulary), both learning systems contribute differentially to vocabulary and syntax, reflecting the presence of both sequential and associative elements in each of these language components (Krishnan, Watkins, & Bishop, 2016). Thus, it is possible that procedural memory will be associated with a greater extent with those aspects of language for which sequential structure is key (e.g., syntax and phonology), but will still be associated-albeit to a lesser extent-with aspects of language that are less intrinsically sequential (e.g., vocabulary). However, to our knowledge, there is as yet no formal set of predictions in the existing literature that specifies the expected relative strengths of these associations.

Several authors have proposed that neurological impairments in the procedural memory system may account for the language and literacy difficulties (i.e., with grammar and phonology) experienced by individuals with DLD and dyslexia, respectively (Fawcett, Nicolson, & Dean, 1996; Krishnan et al., 2016; Nicolson & Fawcett, 2011; Ullman, 2004; Ullman et al., 2020; Ullman & Pullman, 2015). Unlike domain-specific accounts of dyslexia and DLD (e.g., phonological deficit: Snowling, 2000; grammatical deficit: van der Lely, 2005; van der Lely & Pinker, 2014), hypotheses focusing on the deficits in the procedural memory system emerge as an alternative core-deficit account of dyslexia and DLD, aiming to explain the range of profiles observed within these diagnostic categories (Krishnan et al., 2016; Ullman et al., 2020; Ullman & Pierpont, 2005), including nonlinguistic difficulties with motor skills, attention, and working memory (Baird, Dworzynski, Slonims, & Simonoff, 2010; Brookman, McDonald, McDonald, & Bishop, 2013; Buchholz & McKone, 2004; Delage & Frauenfelder, 2020; Fostick & Revah, 2018; Hill, 2001; Romani, Tsouknida, di Betta, & Olson, 2011).

Despite the predicted difficulties in procedural memory, individuals with DLD and dyslexia are thought to have relatively intact declarative memory and thus be able to compensate for their procedural deficits by relying on the declarative system. Thus, while the procedural/declarative model (Lum, Conti-Ramsden, Page, & Ullman, 2012; Ullman & Pierpont, 2005) clearly predicts a positive association between language and literacy and procedural memory in typical populations, the predictions for atypical populations are more complex, given the possible compensatory role of declarative memory. On the one hand, procedural memory and language skills may be correlated within language-disordered groups, even if mean levels of performance are lower than in the typically developing population. On the other hand, if there is active compensation by the declarative system, such that procedural memory is not the main driver of language learning, then a correlation between procedural memory and language would no longer be predicted (Lum et al., 2012). Rather, language and literacy abilities which would usually be expected to rely on the procedural memory system (e.g., grammar and phonology) would instead correlate with declarative memory.

1.2. Review of empirical evidence

The Serial Reaction Time task (SRTT; Nissen & Bullemer, 1987) is by far the most commonly used task in the now substantial body of research examining the procedural learning abilities of typically developing children and adults and those with these developmental disorders. In this task, participants are usually presented with a stimulus that appears in one of four locations on the screen with participants being asked to respond as quickly as possible to its position by pressing the corresponding key in the keyboard. Unbeknownst to the participants, some of the positions of the stimulus follow a pattern. As participants learn the sequence, they begin to implicitly anticipate sequenced trials, resulting in faster response times to these trials compared to random trials (Barker, 2012; Nissen & Bullemer, 1987; Schwarb & Schumacher, 2012). In contrast to other tasks tapping procedural memory, there is consistent evidence of the involvement of the basal ganglia-the core subcortical structure of the procedural learning system-in the SRTT from both neuroimaging (e.g., Janacsek et al., 2020) and patient studies (Williams, 2020). Thus, the SRTT, particularly its “classic” version (Janacsek et al., 2020), is arguably the most well-established task of procedural learning.

Meta-analyses comparing the performance of individuals with dyslexia and DLD to typically developing individuals have found a group deficit in DLD and dyslexia on the SRTT (Lum et al., 2013, 2014; West et al., 2021). The magnitude of the effect size has been found to be significantly moderated by the interaction between the age of participants (Lum et al., 2013, 2014) and the number of exposures to the sequence (Lum, Conti-Ramsden, Morgan, & Ullman, 2014) or the type of sequence (Lum et al., 2013), with smaller effects for older participants when given more practice (Lum et al., 2013) or when these participants were assessed using a second-order conditional sequence (Lum et al., 2014). Yet, as noted by West et al. (2021), the effect size for this group difference tends to be small ($g = -.30$).

Taking an individual differences approach, in which rather than group comparisons, performance on the SRTT is correlated with performance on standardized measures of language, has yielded a more mixed set of findings. In typically developing children, where the prediction is that there should be a positive association, some studies have indeed found correlations between the procedural learning effect captured by the various versions of the SRTT and standardized measures of language (Clark & Lum, 2017a; Desmottes, Meulemans, & Maillart, 2016; Lum et al., 2012). On the other hand, other studies using similar methods have found no association (Gabriel, Meulemans, Parisse, & Maillart, 2015; Henderson & Warmington, 2017; Siegelman & Frost, 2015; Vakil, Lowe, & Goldfus, 2015; West, Clayton, Shanks, & Hulme, 2019). In children with DLD and dyslexia, where the prediction from the Procedural Deficit Hypothesis (Lum et al., 2012) is that the association between language skills and procedural learning should be absent, or at least smaller than in TD, there have also been mixed findings: some studies have found a correlation between procedural learning and language in children with DLD (e.g., Desmottes et al., 2016) and dyslexia (e.g., Vakil et al., 2015), whereas others have not (e.g., DLD: Clark & Lum, 2017; Desmottes, Maillart, et al., 2017; Gabriel et al., 2015; Lum et al., 2012; dyslexia: Deroost et al., 2010; Henderson & Warmington, 2017; West et al., 2019).

Two recent meta-analyses which synthesize this evidence base concluded that there was no convincing evidence of a relationship between procedural learning in the SRTT and continuous measures of language and /literacy measures (Lammertink et al., 2020; West et al., 2021). West et al. (2021) meta-analysis comprehensively examined the role of procedural learning on language and literacy development across a set of tasks considered to tap into procedural learning (e.g., SRTT, Hebb serial order learning task, artificial grammar learning, and statistical learning tasks, among others), though there is a lack of consensus as to whether all these tasks tap into procedural learning (see Conway et al. 2019 for a discussion). The results pertaining to the SRTT are of particular interest to the current study: data from 441 participants (drawn from five studies) revealed a negligible association between procedural learning and measures of language and decoding in typically developing children and adults ($r = .03$; West et al., 2021). Similarly, the large-scale meta-analysis conducted by Lammertink et al. (2020) ($N = 139$ children with DLD and $N = 573$ typically developing children; 19 studies), also found no evidence of an association between procedural learning in the SRTT and expressive (overall sample: $r = .07$; DLD: $r = .03$; TD: $r = .11$) or receptive grammar (overall sample: $r = .05$; DLD: $r = .04$; TD: $r = .06$) in children with or without DLD (Lammertink et al., 2020), with no statistical difference between groups. As discussed above, while the absence of an association between grammar and procedural learning in children with DLD is consistent with the predictions of the procedural deficit hypothesis, this pattern of results for school-aged typically developing children and adults is not.

On the other hand, a contrary set of findings that were more consistent with the predictions of the procedural/declarative model were reported by Hamrick et al. (2018) in a set of four meta-analyses that included 16 studies ($N = 470$ children, $N = 254$ adults) examining the link between language and declarative and procedural memory in first- (children) and second-language learners (adults). In this case, positive associations between procedural learning and grammar were found for children ($r = .27$) and highly proficient second-language adults ($r = .55$), while low-proficiency second-language adults did not show this relationship ($r = -.01$). Conversely, declarative memory was associated with lexical abilities and to a lesser extent with grammar in children ($r = .41$, $r = .16$, respectively), and with grammar in low-proficiency but not high-proficiency second-language adult learners ($r = .47$, $r = -.07$, respectively). These results point to the importance of considering language proficiency as a potential moderator of the relationships between procedural memory and language learning, such that in individuals with low language proficiency (e.g., in second-language learners, or DLD), declarative memory plays a more important role in grammar-learning than procedural memory (Ullman, 2004, 2015, 2016a).

More generally, in interpreting the findings of correlational studies examining procedural learning in relation to language and literacy, it is important to highlight recent studies showing that the SRTT has low reliability both in children and adults (Kalra, Gabrieli, & Finn, 2019; Siegelman & Frost, 2015; Stark-Inbar, Raza, Taylor, & Ivry, 2017; West, Vadillo, Shanks, & Hulme, 2018; West, Shanks, & Hulme, 2021). Poor psychometric properties are particularly problematic for individual difference studies as they contribute to the attenuation of the association between measures (Fleiss, 1986; Rouder et al., 2019; Spearman, 1904). Importantly, this reliability issue is not limited to the SRTT, as other related tasks have been found to have equally poor reliability (e.g., artificial grammar learning: Kalra et al., 2019; Hebb task: Bogaerts, Siegelman, Ben-Porat, & Frost, 2018; West et al., 2018; statistical learning: Arnon, 2019; probabilistic classification task: Kalra et al., 2019).

1.3. *The current study*

In light of these findings, there still seems to be mixed evidence for a relationship between language and procedural learning as indexed by the SRTT. Yet, given the predictions of the procedural deficit hypothesis, which does not propose a unique contribution of procedural learning to the acquisition of grammar, but for all rule-based knowledge, further research on the role of procedural memory on language and literacy development more broadly is required. Thus, contrary to previous meta-analyses which have mostly focused on the relationship between vocabulary and grammar and procedural memory (Hamrick et al., 2018; Lammertink et al., 2020) or analyzed the relationship between language and literacy and procedural memory only in a small sample of studies (West et al., 2021), the present meta-analysis, with the largest sample to date, aims to extend these findings by analyzing the relationship between procedural learning and different subdomains, namely, grammar, vocabulary, phonology, spelling, reading, in school-aged children and adults (5–27 years) with and without language and literacy impairments.

Based on the procedural deficit hypothesis (Ullman & Pullman, 2015), distinct patterns of association are hypothesized to occur for these populations as the declarative memory system has been proposed to compensate for the procedural learning impairments in populations with DLD and dyslexia. Specifically, a stronger association between procedural learning and grammar would be expected for typically developing children than those with DLD (as found by Lum et al., 2012). Similar findings would be expected for children with dyslexia, where the declarative memory system would be expected to compensate for the literacy deficits which are also proposed to emerge as a consequence of procedural learning impairments (Lum et al., 2013; Ullman, 2004; Ullman et al., 2020). Furthermore, since the ability to compensate for language and literacy deficits would be expected to increase with age as the declarative memory system matures (Lum et al., 2013; Ullman, 2004; Ullman & Pullman, 2015), age is expected to be a significant moderator of the association between procedural learning and language and literacy measures, especially for individuals with DLD and dyslexia.

Finally, considering the methodological variability in how the SRTT is designed and implemented across different studies (Schwarb & Schumacher, 2012) and how these methodological decisions have been found to moderate the magnitude of the difference between dyslexic and control groups in the SRTT (Lum et al., 2013, 2014), factors such as the (a) number of trials, (b) type of sequence, and (c) type of SRTT were included as potential moderators. Previous meta-analyses have found that studies with a larger number of trials (Lum et al., 2013, 2014) showed a smaller difference between disordered and typically developing groups. Similarly, the size of the effect was also found to be moderated by the type of sequence, with second-order conditional sequences being associated with smaller effect sizes when comparing children with and without dyslexia in the SRTT (Lum et al., 2013), and type of SRTT, with deterministic sequences showing larger effect sizes than alternating sequences (West et al., 2021) although note that this was not the case in Lammertink et al. (2020), who did not find a moderating effect for the type of sequence. Thus, by taking these factors into account, we aim to determine whether methodological decisions impact the strength of the association between language and literacy and procedural learning tasks and thus have contributed to the inconsistent pattern of findings across studies.

1.4. Aims and research questions

The current meta-analysis aims to provide a more comprehensive analysis of the relationship between language and literacy and procedural memory. By including language and literacy measures across subdomains, this meta-analysis will be able to assess the predictions of the procedural deficit hypothesis in children and adults with and without language and literacy impairments. Furthermore, by investigating the effect of the moderating variables, this meta-analysis may also provide some possible explanations for the inconsistent pattern of results in the literature. Unlike previous meta-analyses, the present study includes effect sizes for all available language and literacy measures as, instead of aggregating the effect sizes derived from a single sample, we take advantage of multilevel models to deal with the nonindependent effect sizes, thus preventing information loss.

We addressed the following preregistered (https://osf.io/vtdg3) research questions and hypotheses:

Research Question 1: We examined the relationship between procedural learning and language/literacy abilities in typically developing adults and children. Specifically, following the declarative/procedural model, we predicted correlations between procedural learning and grammar, phonology, reading, and spelling (Ullman et al., 2020), while vocabulary was expected to only be weakly correlated with procedural learning (Hamrick et al., 2018; Ullman, 2004; Ullman et al., 2020) (Hypothesis 1 (H1).

Research Question 2: In line with the procedural deficit hypothesis (Ullman, 2004; Ullman et al., 2020; Ullman & Pullman, 2015), we anticipated that group membership would moderate the relationship between procedural learning and language/literacy abilities, with (a) stronger associations expected between grammar, phonology, and procedural learning for typically developing groups than DLD groups, based on the proposal that individuals with DLD compensate for difficulties in grammatical and phonological acquisition with the declarative learning system (H2) (Ullman, 2004; Ullman & Pullman, 2015; Ullman et al., 2020); and similarly (b) stronger correlations between phonology, reading, spelling, and procedural learning for typically developing groups than for dyslexic groups, as individuals with developmental dyslexia would be predicted to compensate for phonology, reading, and spelling difficulties with the declarative learning system (H3) (Ullman, 2004; Ullman & Pullman, 2015; Ullman et al., 2020).

In a set of exploratory analyses, we also examined whether age, the number of SRTT sessions, sequence complexity, and SRTT type (deterministic vs. probabilistic sequences) moderate the relationship between procedural learning and language/literacy abilities.

2. Methods

All materials for this meta-analysis are available (https://osf.io/ev2xw/), including the dataset and scripts necessary to replicate all reported analyses and plotting.

2.1. Search strategy

To find eligible studies, literature searches were conducted up to November of 2020 on Pubmed and Google Scholar using the following search terms:

> PUBMED: procedural learning OR procedural memory OR sequence learning OR implicit learning AND language OR reading OR dyslexia OR language impairments
>
> procedural learning OR procedural memory OR sequence learning OR implicit learning AND SRT OR Serial Reaction Time task AND language OR reading OR dyslexia OR language impairments
>
> GOOGLE SCHOLAR: Relationship AND language AND "serial reaction time task" AND visual sequence learning

Once the search was completed, the first author screened all titles and abstracts (Fig. 1) for records which analyzed the procedural learning, language, or literacy skills of individuals with and without dyslexia and developmental language disorder. Sixty-seven records met these criteria and were then subjected to a full-text analysis against the inclusion and exclusion criteria to determine eligibility. Full-text eligibility was assessed using the following inclusion criteria: (i) Population: TD controls and/or individuals with language/reading impairments of speakers of alphabetic languages; (ii) Used a strictly visual deterministic, probabilistic, or alternating SRTT with procedural learning computed as the difference between sequenced/probable and random/improbable trials; Audio-visual SRTTs or tasks that alternated types of statistical dependencies were not included (e.g., Jackson, Leitão, Claessen, & Boyes, 2020); (iii) Analyzed the relationship between language/literacy and procedural learning and reported Pearson's correlation (or equivalent) coefficients; (iv) If correlations were missing but the required measures were used, we solicited the correlation coefficients directly from the authors; (v) English language publication; (vi) If the same results were published in multiple articles, these were only reported once in the meta-analysis; we selected the

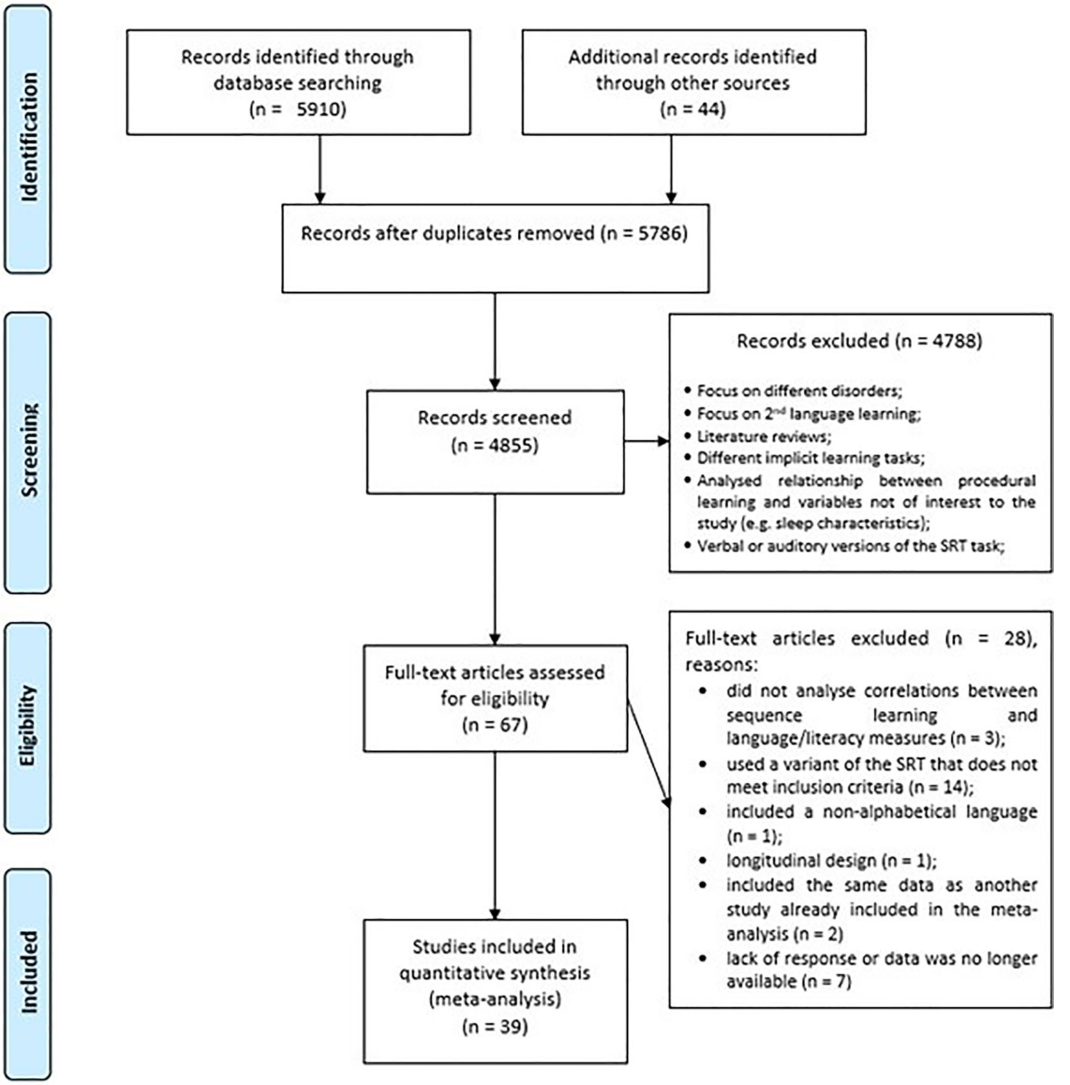

Fig. 1. PRISMA flowchart showing a selection of studies for meta-analysis on the relationship between language and literacy and procedural learning.

publication with the largest sample size and most comprehensive information; (vii) Publication dates: between 2000 and 2020 to increase the chances of data availability; and exclusion criteria: (i) Second language learning studies (we opted to exclude second language learning studies as the predictions from the procedural/declarative model are distinct for first and second language learners (e.g., Hamrick et al., 2018); dual task paradigms with the SRTT; longitudinal studies that measured procedural learning and language measures at different time points (e.g., language measures in infancy and procedural learning in adulthood); SRTTs which only involved three positions as typically four positions are used; studies that used adaptations that deviate substantially from the task described by Nissen and Bullemer (1987) (e.g., in the task used by Vicari, Marotta, Menghini, Molinari, and Petrosini, 2003, participants were expected to only provide a motor response to a subset of the stimuli).

Forward and backward searches were conducted on these records. Discrepancies regarding an article's eligibility were resolved among the authors until a consensus was reached. Once the list of articles to be included in the meta-analysis was agreed upon, it was sent to two independent experts in the field for feedback to ensure that relevant articles were not missed.

2.2. *Data extraction*

Articles included in the meta-analysis were coded by the first author and a second coder blind to the purpose of the study using the preregistered data extraction form developed for the current review and available at https://osf.io/ev2xw/. Data extraction was compared until 100% agreement was reached between coders. The following moderators were included: age of participants, group (TD/typically developing, DD/dyslexic, or DLD/development language disordered groups), domain of interest (language or literacy), subdomains (phonology, vocabulary, grammar, reading, or spelling[1]), type of SRTT (deterministic, probabilistic, or alternating), sequence complexity (first- [FOC] or second-order conditional [SOC]), number of trials the participants were exposed to, and number of sessions (including training and testing sessions).

2.3. *Meta-analytic approach*

The effect size metric of Pearson's *r* was used to represent the strength and direction of the relationship between procedural learning and language/literacy skills. Correlation coefficients were recoded to reflect a positive relationship (higher procedural learning and higher language/literacy skills) when necessary. When missing data were identified, the corresponding authors were contacted by email requesting the relevant information; this had the aim of ensuring all correlations-rather than only significant ones reported in the published papers-were included, as reporting biases of this nature have been shown to inflate the overall effect (Kirkham et al., 2010). In the absence of a reply or data being unavailable, the articles were not included in the meta-analysis.

All correlation coefficients were converted from Pearson's *r* to Fisher's *z* scale as Pearson's *r* is not normally distributed (Hedges & Olkin, 1985). Since most studies contributed multiple correlation coefficients to the meta-analysis, to deal with the lack of independence across effect sizes and avoid reducing power by calculating the average for the effect sizes for these studies, robust variance estimation (RVE) was used for model estimation alongside small-sample corrections (Hedges, Tipton, & Johnson, 2010; Tipton & Pustejovsky, 2015) via the *robumeta* package for R (Fisher & Tipton, 2015). RVE methods use a working model that approximates the dependence structure but does not require exact knowledge of the error distribution or covariance structure between effect size estimates. By using RVE methods, even when the working model is misspecified, the meta-regression coefficient estimates will be unbiased (Fisher & Tipton, 2015; J. E. Pustejovsky & Tipton, 2022).

2.4. Modeling

An intercept-only meta-regression was initially run to estimate the overall effect size between procedural learning and language/literacy across groups. Two separate random effects models for language and literacy with correlated effects dependencies were computed using the *robumeta* package (Fisher & Tipton, 2015) to deal with the nonindependent effect sizes. Even though there is also evidence of hierarchical dependencies (e.g., some research groups contributed multiple studies to the meta-analysis), weights for correlational dependency were selected as recommended by Tanner-Smith, Tipton, and Polanin (2016) since this was the most common form of dependency in the present dataset as a large number of studies contributed more than one effect size per sample. For all analyses, the in-study effect size correlation (ρ) was set at .8. In addition, sensitivity analyses were performed across varying values of rho (.0, .2, .4, .6, .8, 1.0) to assess whether results were robust to changes in rho values.

To answer the research questions, separate RVE mixed-effects meta-regression models were performed with *group* as a moderating variable to determine whether there is a significant correlation between language and literacy for the typically developing group and disordered groups (represented in Table 2 as "$TD = DLD = DD = 0$," with DD standing for the dyslexic group) and whether the magnitude of this effect is greater for this group than for the clinical groups (represented in Table 2 as "*TD vs. DLD vs. DD*" or a subset of the groups depending on the research questions). Separate meta-regressions were also conducted for each of the remaining moderator variables (e.g., age, type of sequence) to assess their effect on the association between procedural learning and language/literacy for the overall sample (Tables 2 and 4), with the categorical levels contrasted in a similar manner to group (e.g., "FOC vs. SOC"-representing the contrast between first- and second-order conditional sequences) using the *Wald_test()* function from the *clubSandwich* package (Pustejovsky, 2021). This function uses a method called approximate Hotelling's T^2 test which has been shown to perform adequately even with degrees of freedom close to 0 (Tipton & Pustejovsky, 2015).

After performing the meta-analytic calculations, Fisher's z overall estimates were converted back to Pearson's r for reporting the average correlation and 95% confidence interval (CI) for each model.

2.5. Bias and heterogeneity analyses

To test for the presence of publication/reporting bias, while taking account of the dependencies in the data analysis, the PET-PEESE estimates (Stanley & Doucouliagos, 2014) were computed using RVE via the *robumeta* package (Fisher & Tipton, 2015). PET (Precision-Effect Test) and PEESE (Precision-Effect Estimate with Standard Error) methods use the standard error and the sampling variance as predictors, respectively. The two-step method PET-PEESE is recommended because both PET and PEESE methods show bias, with PET being downwardly biased when the true effect is different from zero, while PEESE shows an upward bias when the true effect is zero (Stanley & Doucouliagos, 2014). The two-step process involves first assessing whether the PET estimate is significant, if the effect is significant, then the PEESE estimate is used; if not, the PET estimate should be adopted (Stanley & Doucouliagos, 2014).

Since not all classical methods are available to analyze outliers, influential cases, and bias for models with dependency, effect sizes were aggregated via the *agg* function from the *MAc* package (Del Re & Hoyt, 2018). Following Borestein (2009), the default correlation between within-study effect sizes was set at .50, with complementary sensitivity analyses to ensure the robustness of the results ($r = .10, .30, .50, .70, .90$). The *influence* function from the *metafor* package was used to identify potential outliers and influential cases for the aggregated effect sizes for each model (Viechtbauer, 2010). To detect evidence of publication bias, funnel and contour-enhanced funnel plots were also produced (Galbraith, 1988; Light et al., 1984) using the *metafor* package (Viechtbauer, 2010). Rank correlation tests were performed to assess funnel plot asymmetry (Begg & Mazumdar, 1994).

Study heterogeneity was analyzed using the Q statistics, I^2, and τ^2 (Higgins & Thompson, 2002). The Q test assesses heterogeneity by comparing the effect sizes across studies to determine whether all studies show the same effect (null hypothesis). However, this test has been shown to be poor at detecting true heterogeneity in smaller samples due to a lack of power. Thus, the results of the Q test need to be considered in light of other measures of heterogeneity. I^2 represents the proportion of variability across studies due to heterogeneity rather than chance. This measure was introduced by Higgins and Thompson (2002) as being more interpretable and comparable across studies than Q statistics. Yet, given the reliance on the Q statistics for its calculation, I^2 is also often imprecise and/or biased in small samples (von Hippel, 2015). For ease of interpretation, heterogeneity above 50% and 75% tends to be considered moderate and substantial, respectively (Higgins & Thompson, 2002). Finally, τ^2-the variance in the true effect sizes-was also interpreted since random effects models were adopted across analyses as we anticipated that studies did not represent a homogeneous population due to the methodological and sampling differences (Borestein, 2009).

2.6. Correction for attenuation

In classical test theory, observed scores are thought to reflect the true scores plus measurement error (Novick, 1966), with measurement error often limiting the size of the correlation between two variables (Spearman, 1904; Wiernik & Dahlke, 2019). Thus, as recommended by Wiernik and Dahlke (2019), the pooled correlation coefficient between language and literacy and procedural learning was corrected for attenuation using Spearman's derivation (Spearman, 1904).

$$true\ correlation = \frac{observed\ correlation}{\sqrt{reliability\,(x)\,.\,reliability\,(y)}}$$

This method adjusts the raw correlation by taking into consideration the reliability estimates for each measure (Spearman, 1904). Considering the scarce information about measurement reliability, an artifact distribution method was used, as measurement error correction was only performed after model estimation (Wiernik & Dahlke, 2019).

The overall reliability for the SRTT was estimated to be approximately .30, based on a recent meta-analysis conducted by Oliveira, Hayiou-Thomas et al. (2023, April 25). Unfortunately, only a small number of studies reported the reliability of the tasks adopted (namely, Siegelman & Frost, 2015; West et al., 2018; West et al., 2021), thus the estimate may not be representative of the reliability of the SRTT used in the studies included in this meta-analysis. For the literacy and language measures, an overall reliability of .70 was selected based on the frequent test-retest reliability of standardized measures used to assess language and literacy. However, as discussed by Rouder et al. (2019, March 25), disattenuation methods are flawed and can produce highly variable estimates which can be both inflated and deflated, so these need to be interpreted with caution.

3. Results

In total, the meta-analysis comprised 39 independent studies (see Table 1 for more details), summarizing 500 effect sizes and data from 2396 participants. Participants' age range: [5.2, 27.7], $M = 12.69$, $SD = 5.64$. See https://osf.io/ev2xw/ for the entire dataset.

3.1. Relationship between language and literacy and procedural learning

To directly test the predictions of the procedural/declarative model and the procedural deficit hypothesis, separate analyses were required for the typically developing, DLD, and dyslexic groups. Yet, before presenting those results, we started by examining the overall effects across populations for both literacy and language.

3.1.1. Overall effect

All studies ($k = 39$) were included in this analysis (citations marked with an asterisk). The estimated average correlation between procedural learning and overall language and literacy measures for all studies using RVE was Fisher's $z = .06$, 95% CI [.007, .12], SE = .03, $t(32.7) = 2.3$, $p = .028$, with a Pearson's r of .06, indicating a very modest-though statistically significant at this sample size-association between procedural learning and language/literacy. Sensitivity analyses indicated that the results are robust to different values of rho (Fisher's z varied between .0619 and .0620). There was no significant difference between language and literacy measures in terms of the strength of the relationship with procedural learning ($F(1, 22.3) = .34$, $p = .563$). Note that this result is not necessarily at odds with our predictions as the relationship between language and literacy and procedural memory may have been shifted downwardly by the inclusion of effect sizes from the disordered groups. A direct test of our predictions, based on the procedural/declarative model, requires separate analyses for TD and disordered groups.

To further explore the relationship between procedural learning and language/literacy, separate RVE models were computed for language and literacy measures.

Table 1
Overview of the study sample characteristics for each individual experiment in our sample

Study	Group	*N*	Mean age (years)	Type of SRTT	Sequence complexity
Clark and Lum (2017a)	DLD	20	8.92	deterministic	FOC
	TD	20	9.09	deterministic	FOC
Clark and Lum (2017b)	DLD	25	9.81	deterministic	FOC
	TD	27	9.70	deterministic	FOC
Deroost et al. (2010)	DD	28	13.50	deterministic	FOC
Desmottes et al. (2016b)	DLD	21	10.20	deterministic	FOC
	TD	21	10.54	deterministic	FOC
Desmottes, Maillart, and Meulemans	DLD	18	10.24	deterministic	FOC
(2017a)	TD	17	10.12	deterministic	FOC
Desmottes, Meulemans, Patinec, and	DLD	15	10.39	deterministic	FOC
Maillart (2017b)	TD	15	10.41	deterministic	FOC
Earle and Ullman (2021)	DLD	21	20.52	deterministic	
	TD	79	20.49	deterministic	
Gabay, Schiff, and Vakil (2012)	DD	12	23.58	deterministic	SOC
	TD	12	24.83	deterministic	SOC
Gabriel, Maillart, Guillaume, Stefaniak,	DLD	15	10.17	probabilistic	SOC
and Meulemans (2011)	TD	15	10.25	probabilistic	SOC
Gabriel, Stefaniak, Maillart, Schmitz,	DLD	15	10.25	deterministic	FOC
and Meulemans (2012)	TD	15	10.42	deterministic	FOC
Gabriel et al. (2013)	DLD	23	9.67	deterministic	SOC
	TD	23	9.58	deterministic	SOC
Gabriel et al. (2015)	DLD	16	9.92	deterministic	FOC
	TD	16	9.83	deterministic	FOC
Hedenius et al. (2013)	DD	12	11.00	alternating	SOC
	TD	17	11.10	alternating	SOC
Hedenius et al. (2021)	DD	30	11.60	alternating	SOC
	TD	32	11.68	alternating	SOC
Henderson and Warmington (2017)	DD	30	21.13	deterministic	SOC
	TD	29	20.31	deterministic	SOC
Hsu and Bishop (2014)	DLD	48	8.80	deterministic	FOC
	TD	24	7.37	deterministic	FOC
Kidd and Kirjavainen (2011)	TD	120	5.20	deterministic	FOC

(*Continued*)

Table 1
(Continued)

Study	Group	*N*	Mean age (years)	Type of SRTT	Sequence complexity
Kidd (2012)	TD	100	5.58	deterministic	FOC
Kuppuraj, Rao, and Bishop (2016)	DLD	30	11.65	deterministic	SOC
	TD	30	10.68	deterministic	SOC
Lammertink et al. (2020)	DLD	35	9.08	deterministic	FOC
	TD	35	9.08	deterministic	FOC
Lee and Tomblin (2015)	DLD	25	22.14	deterministic	SOC
	TD	23	22.23	deterministic	SOC
Llompart and Dąbrowska (2020)	TD	60	21.63	deterministic	SOC
Lukacs and Kemeny (2014)	DLD	29	9.10	deterministic	FOC
Lum and Kidd (2012)	TD	58	5.45	deterministic	FOC
Lum et al. (2012)	DLD	51	9.80	deterministic	SOC
	TD	51	9.85	deterministic	SOC
Mayor-Dubois, Zesiger, Van der Linden, and Roulet-Perez (2014)	DLD	65	18.00	deterministic	FOC
Menghini et al. (2010)	DD	60	11.43	deterministic	FOC
	TD	65	11.94	deterministic	FOC
Mimeau, Coleman, and Donlan (2016)	TD	76	6.50	deterministic	FOC
Oliveira et al. (2022, May 10)	TD	47	20.09	probabilistic	SOC
Park et al. (2018)	DLD	27	9.95	deterministic	FOC
	TD	59	10.27	deterministic	FOC
Schmalz, Moll, Mulatti, and Schulte-Körne (2019)	TD	84	27.70	deterministic	
Siegelman and Frost (2015)	TD	76	24.10	probabilistic	SOC
Spit and Rispens (2019)	TD	25	10.90	deterministic	
Stoodley, Harrison, and Stein (2006)	DD	19	23.92	deterministic	SOC
	TD	21	22.83	deterministic	SOC
Vakil et al. (2015)	DD	23	12.41	deterministic	SOC
	TD	30	12.55	deterministic	SOC
van Witteloostuijn, Boersma, Wijnen, and Rispens (2019)	DD	50	9.83	deterministic	FOC
	TD	50	9.67	deterministic	FOC
West et al. (2018)	TD	101	8.19	probabilistic	SOC
West et al. (2019)	DD	23	9.81	probabilistic	SOC
	TD	25	7.62	probabilistic	SOC
West et al. (2021)	TD	112	7.68	probabilistic	SOC

Abbreviations: DD, developmental dyslexia; DLD, development language disorder; *N*, number of participants; TD, typically developing children.

3.1.2. Language and procedural learning

3.1.2.1. Overall effect across participants: Thirty-five studies reported the relationship between procedural learning and language (marked with an asterisk). Fisher r-to-z transformed correlation coefficients ranged from −.97 to .91, with just over half of the estimates being positive (55%). However, the average correlation between procedural learning and spoken language measures was again very modest, though statistically significant: Fisher's $z =$ .06, 95% CI [−.002, .12] SE = .03, $t(30.1) = 1.98$, $p = .057$, with an equivalent Pearson's correlation of $r = .06$, 95% CI [−.002, .12]. Sensitivity analyses confirmed that this result was robust to different levels of rho. There was a moderate level of heterogeneity in the effect sizes $\tau^2 = .028$, $I^2 = 51.72$, which was further explored using meta-regression analyses (presented in Tables 2 and 3). As before, this result offers no support for or against our predictions given that the overall estimate may be smaller due to the inclusion of the disordered groups.

3.1.2.2. Moderator analyses: Results of all separate RVE meta-regressions with moderator variables as predictors are shown in Tables 2 and 3. F-tests were used to compare the estimates between levels of categorical predictors.

There was no evidence that group membership affected the magnitude of the pooled association between procedural learning and language (Table 2). Subgroup RVE meta-analyses (Table 3) for each linguistic domain (grammar, vocabulary, and phonology) were conducted separately to assess whether there were group differences in the relationship between language subdomains and procedural learning (H1/H2/H3). Since only one study included vocabulary measures for the dyslexic group, this group was removed from the vocabulary meta-regressions. The overall pattern was consistent across analyses, with no evidence of a significant relationship between procedural learning and language subdomains regardless of group (grammar: $F(3, 3.02) = 1.40, p = .394$; phonology: $F(3, 67.83) = .78, p = .539$; vocabulary: $F(2, 11.5) = 1.3$, $p = .311$) and there was no evidence that the strength of the association differed between the typically developing and DLD groups for grammar ($F(1, 17.1)$ $= 1.07$, $p = .315$) or phonology specifically ($F(1, 9.02) = 2.44$, $p = .153$), nor between the typically developing and dyslexic groups for phonology ($F(1, 7.95) = 0.628, p = .451$). Thus, there was no evidence supporting H1 as there was no evidence of a relationship between language and procedural memory on the SRTT for the typically developing group. Furthermore, contrary to our predictions, the relationship between procedural learning and phonology did not differ between the typically developing and disordered groups (H2, H3), nor did the association between grammar and procedural memory differ between the typically developing and the DLD group (H2). Disattenuated correlations for the group comparisons (represented as R) are also presented in Table 3. These show small correlations for the typically developing group between procedural learning and grammar ($R = .12$) and vocabulary ($R = .19$). Small associations between procedural learning and phonology ($R = .10$) and grammar ($R =$ .19) were also observed for the dyslexic group. For the DLD, on the other hand, there was a moderate association between phonology and procedural learning ($R = .46$). All the other correlations were negligible ($R < .1$).

The moderating effect of sampling and methodological differences (domain, age of participants, sequence complexity, type of SRTT, session, or number of trials) was also tested on

Table 2
Results of all separate meta-regressions with moderator variables for language measures

Moderator (bolded) and level	Study characteristics			Effect size		Test of significance					Heterogeneity		Sensitivity analysis
	s	*k*	*F*	Fisher's *z*	*r*	SE	*t*	*p*	95% CI		τ^2	I^2	range
Group	35	323	—	—	—	—	—	—	—	—	.029	52.572	—
TD = DLD = DD = 0	—	—	1.23	—	—	—	—	.333	—	—	—	—	—
TD vs. DLD vs. DD	—	—	.124	—	—	—	—	.884	—	—	—	—	—
TD	—	—	—	.056	.056	.033	1.719	.100	−.012	.125	—	—	.0564; .0565
DLD	—	—	—	.083	.082	.090	.917	.376	−.112	.278	—	—	.0826; .0830
DD	—	—	—	.030	.030	.060	.500	.637	−.120	.179	—	—	.0298; .0299
Subdomain	35	313	—	—	—	—	—	–	—	—	.030	53.122	—
Grammar = Phonology = Vocabulary = 0	—	—	1.26	—	—	—	—	.318	—	—	—	—	—
Grammar vs. Phonology vs. Vocabulary	—	—	.213	—	—	—	—	.810	—	—	—	—	—
Grammar	—	—	—	.041	.041	.039	1.050	.307	−.041	.122	—	—	.0405; .0407
Phonology	—	—	—	.088	.088	.079	1.120	.291	−.089	.266	—	—	.0883; .0886
Vocabulary	—	—	—	.067	.067	.044	1.540	.160	−.032	.166	—	—	.0669; .0672
Age	35	323	—	.003	.003	.006	.575	.578	−.009	.015	.029	52.610	.0226; .0227
Sequence complexity	33	298	—	—	—	—	—	—	—	—	.026	50.631	—
FOC = SOC = 0	—	—	2.85	—	—	—	—	.082	—	—	—	—	—
FOC vs. SOC	—	—	.130	—	—	—	—	.721	—	—	—	—	—
FOC	—	—	—	.081	.081	.045	1.810	.091	−.015	.177	—	—	.0813; .0815
SOC	—	—	—	.060	.060	.037	1.650	.124	−.019	.140	—	—	.0603; .0606
Session[a]	33	291	—	—	—	—	—	—	—	—	.028	52.786	—
T1 = T2 = T3 = 0	—	—	1.02	—	—	—	—	.462	—	—	—	—	—
T1 vs. T2 vs. T3	—	—	.930	—	—	—	—	.485	—	—	—	—	—
T1	—	—	—	.057	.057	.035	1.629	.115	−.015	.130	—	—	.0574; .0576
T2	—	—	—	.091	.090	.076	1.193	.282	−.099	.280	—	—	.0904; .0908
T3	—	—	—	−.02	−.02	.048	−.397	.732	−.241	.203	—	—	−.0189
# of Trials	35	323	—	—	—	.000	.062	.953	.000	.000	.029	52.591	—
Type of SRTT	35	323	—	—	—	—	—	—	—	—	.031	53.977	—
Det = Prob = Alt = 0	—	—	3.09	—	—	—	—	.213	—	—	—	—	—
Det vs. Prob vs. Alt	—	—	.494	—	—	—	—	.677	—	—	—	—	—
Det	—	—	—	.067	.067	.038	1.788	.086	−.010	.145	—	—	.0673; .0677
Prob	—	—	—	.035	.035	.053	.661	.549	−.120	.190	—	—	.0351; .0357
Alt	—	—	—	.021	.021	.006	3.596	.173	−.120	.190	—	—	.0206; .0207

Notes. *F* values are from Approximate Hotelling–Zhang with small sample correction omnibus tests of the effects of moderators with more than two levels; *r* = Pearson's *R* correlation; standard errors (SE) and *t* values for individual levels of a moderator; *p* values correspond to *F* or *t* values; 95% CI corresponds to the Fisher's *z*.

Abbreviations: *k*, number of effect size estimates; *s*, number of studies.

[a]Studies by Clark and Lum (2017b) and Desmottes et al. (2017) were removed from the meta-regression with session as a moderator variable as these studies did not compute correlations independently for each session.

Table 3
Results of all meta-regressions for language and literacy components with group as a moderator variable

Moderators (bolded) and levels	Study characteristics			Effect size			Test of significance					Heterogeneity		Sensitivity analysis
	s	*k*	*F*	Fisher's *z*	*r*	*R*	SE	*t*	*p*	95% CI		τ^2	I^2	range
Grammar	25	137	—	—	—	—	—	—	—	—	—	0.023	44.432	—
TD = DLD = DD = 0	—	—	1.400	—	—	—	—	—	.394	—	—	—	—	—
TD vs. DLD	—	—	0.210	—	—	—	—	—	.722	—	—	—	—	—
GroupDD	—	—	—	0.085	.085	.185	0.039	2.180	.274	−0.410	0.580	—	—	0.0849; 0.0851
GroupDLD	—	—	—	−0.017	−.017	−.037	0.060	−0.283	.782	−0.147	0.113	—	—	−0.0169; −0.0168
GroupTD	—	—	—	0.057	.057	.125	0.044	1.298	.214	−0.037	0.151	—	—	0.0572; 0.0574
Vocabulary[a]	18	63	—	—	—	—	—	—	—	—	—	0.017	35.482	—
TD = DLD = 0	—	—	1.300	—	—	—	—	—	.311	—	—	—	—	—
TD vs. DLD	—	—	1.490	—	—	—	—	—	.243	—	—	—	—	—
GroupDLD	—	—	—	0.014	.014	.031	0.036	0.389	.706	−0.067	0.096	—	—	0.0141; 0.0142
GroupTD	—	—	—	0.089	.089	.194	0.054	1.665	.129	−0.031	0.210	—	—	0.0892; 0.0895
Phonology	16	111	—	—	—	—	—	—	—	—	—	0.028	42.235	—
TD = DLD = DD = 0	—	—	0.778	—	—	—	—	—	.539	—	—	—	—	—
TD vs. DLD	—	—	2.440	—	—	—	—	—	.153	—	—	—	—	—
TD vs. DD	—	—	0.628	—	—	—	—	—	.451	—	—	—	—	—
GroupDD	—	—	—	0.047	.047	.103	0.065	0.725	.498	−0.116	0.211	—	—	—
GroupDLD	—	—	—	0.212	.209	.455	0.138	1.536	.192	−0.155	0.579	—	—	—
GroupTD	—	—	—	−0.018	−.017	−.038	0.041	−0.427	.678	−0.109	0.074	—	—	—
Reading	15	132	—	—	—	—	—	—	—	—	—	0.003	10.643	—
TD = DLD = DD = 0	—	—	2.120	—	—	—	—	—	.207	—	—	—	—	—
TD vs. DD	—	—	0.003	—	—	—	—	—	.986	—	—	—	—	—
GroupDD	—	—	—	0.045	.045	0.098	0.075	0.596	.576	−0.145	0.235	—	—	0.0442; 0.0477
GroupDLD	—	—	—	0.100[c]	.100	0.217	0.043	2.307	.114	−0.047	0.246	—	—	0.9926; 0.0999
GroupTD	—	—	—	0.047	.047	0.102	0.040	1.167	.283	−0.049	0.142	—	—	0.0462; 0.0488
Spelling[b]	5	12	—	—	—	—	—	—	—	—	—	0.017	34.435	—
TD = DD = 0	—	—	1.600	—	—	—	—	—	.388	—	—	—	—	—
TD vs. DD	—	—	0.896	—	—	—	—	—	.421	—	—	—	—	—
GroupDD	—	—	—	0.252	.247	.538	0.159	1.588	.263	−0.489	0.993	—	—	0.2517; 0.2527
GroupTD	—	—	—	0.047[c]	.047	.103	0.088	0.537	.627	−0.226	0.320	—	—	0.0473; 0.0470

Notes. *F* values are from Approximate Hotelling–Zhang with small sample correction omnibus tests of the effects of moderators with more than two levels; *r* = Pearson's *R* correlation; *R* = disattenuated Pearson's *R* correlation; standard errors (SE) and *t* values for individual levels of a moderator; *p* values correspond to *F* or *t* values; 95% CI corresponds to the Fisher's *z*.

Abbreviations: *k*, number of effect size estimates; *s*, number of studies.

[a]The DD ($k = 2$) and [b]DLD ($k = 1$) groups were not included in these analyses due to the small sample size.

the entire sample, yet there was no evidence of a moderating effect of any of these factors on the relationship between language and procedural memory (Table 2).

3.1.2.3. Model diagnostics: Sensitivity analyses revealed consistent results across rho values for all meta-analytic analyses (see Tables 2 and 3). To assess the presence of influential studies, effect sizes were aggregated per study as these analyses are not available for RVE models. There was no evidence of influential studies.

3.1.2.4. Publication bias: Several assessments of publication bias were conducted on the aggregated data, with the exception of PET and PEESE models which examined publication bias on all effect sizes. Funnel plot and contour plots for the estimates are shown in Figs. 2a and b. Visual inspection shows no obvious evidence of plot asymmetry or overrepresentation of studies in the significance contours. These results are further supported by the nonsignificance of the rank correlation test (Kendall's $\tau = -.03$, $p = .880$).

In line with previous results, both PET and PEESE models with RVE showed no evidence of publication bias (PETrve: $b1 = -.739$, $p = .120$; PEESErve: $b1 = -1.560$, $p = .210$).

3.1.3. Literacy and procedural learning

3.1.3.1. Overall effect: A total of $k = 18$ studies were included in the analyses. Similar findings to those obtained for language were observed for literacy, both for the overall effect, and the meta-regressions examining moderators. Fisher's z for individual studies varied between −.48 and .91, with 57% positive effect sizes. The meta-analytical model revealed a significant, but again very modest, relationship between literacy and procedural learning, Fisher's $z = .05$ (Pearson's $r = .05$), 95% CI [.003, .10], SE = .02, $t(13.5) = 2.29$, $p = .039$, with evidence of a small amount of heterogeneity in the effect sizes ($\tau^2 = .011$, $I^2 = 29.48$). Sensitivity analyses indicated that this result was consistent across rho values (values ranged from .0495 to .0496). Similarly to previous results, the weak relationship between literacy and procedural learning may reflect the inclusion of the disorder groups, thus it does not speak directly to our hypotheses.

3.1.3.2. Moderator analyses: All results from the meta-regressions are presented in Tables 3 and 4.

Group membership was not a significant predictor of the magnitude of the relationship between literacy and procedural memory and there was no evidence that this relationship differed from zero for all groups (Table 4). To answer our research questions (H1/H2/H3), separate meta-regressions were conducted for reading and spelling to determine whether there were differences between groups for these subdomains. Given the low number of effect sizes for the DLD group for spelling ($n = 1$), the DLD group was omitted from these analyses, but was included in the remaining analyses. There was no evidence that the magnitude of the pooled effect size differed between the typically developing and dyslexic groups for spelling ($F(1, 2.7) = .90$, $p = .421$) and reading ($F(1, 7.44) = .003$, $p = .956$). Additionally, the relationship between procedural learning and spelling was not significant for any group ($F(2, 1.95) = 1.60$, $p = .388$). The same pattern was observed for reading ($F(2, 7.64) = .99$,

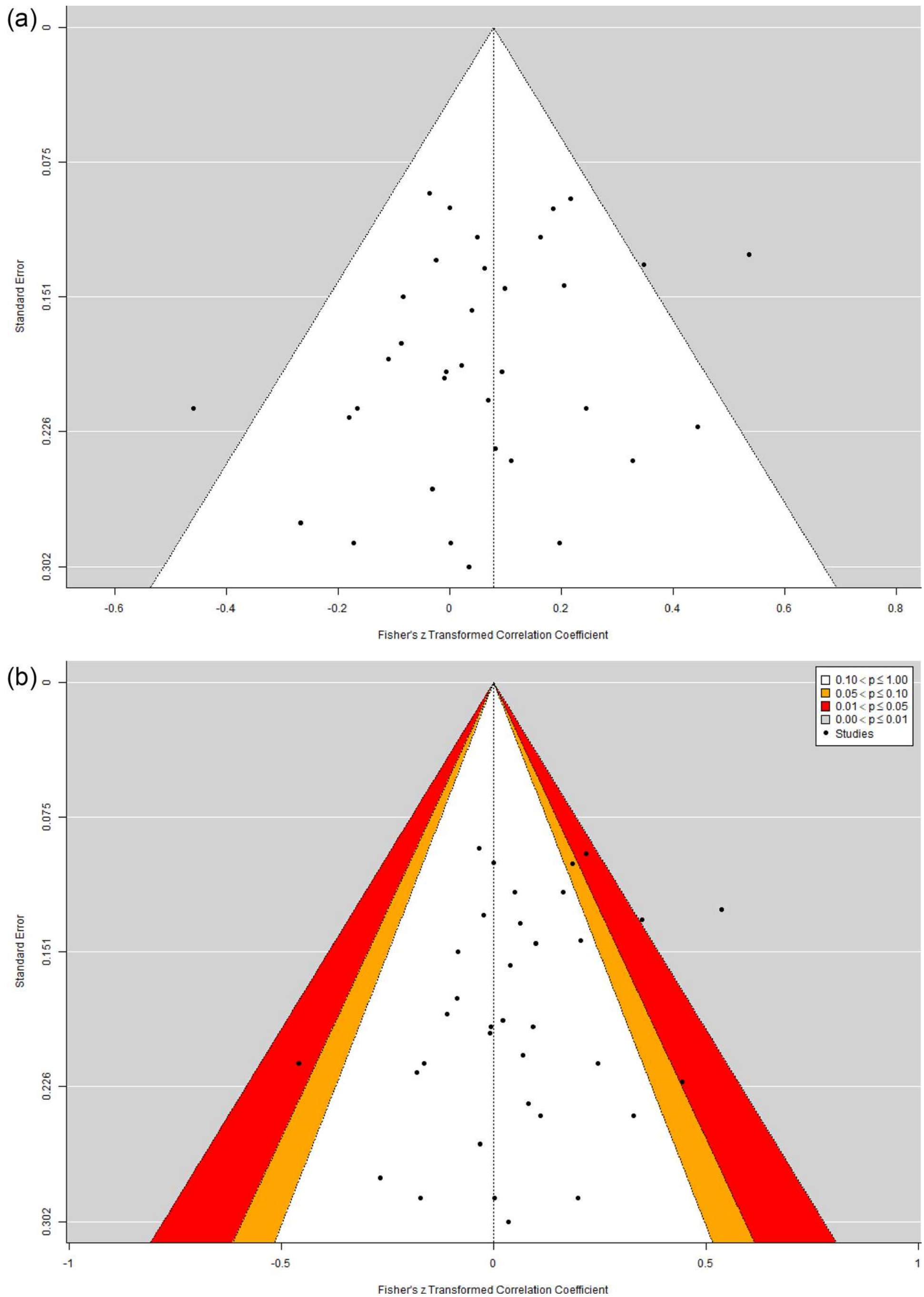

Fig. 2. Funnel plot showing study-level effect sizes plotted against standard error. An asymmetric distribution is taken as evidence of publication bias; (a) funnel plot (left panel) and (b) contour-enhanced funnel plot (right panel).

Table 4
Results of all separate meta-regressions with moderator variables for literacy measures

Moderator (bolded) and level	Study characteristics			Effect size		Test of significance					Heterogeneity		Sensitivity analysis
	s	k	F	Fisher's z	r	SE	t	p	95% CI		τ^2	I^2	range
Group	18	155	—	—	—	—	—	—	—	—	.013	32.056	—
TD = DLD = DD = 0	—	—	1.300	—	—	—	—	.347	—	—	—	—	—
TD vs. DLD vs. DD	—	—	0.127	—	—	—	—	.883	—	—	—	—	—
TD	—	—	—	.038	.037	.030	1.230	.249	−.031	.106	—	—	.0801; .0808
DLD	—	—	—	.054	.054	.069	0.786	.493	−.174	.281	—	—	.0532; .0539
DD	—	—	—	.080	.080	.065	1.237	.254	−.072	.232	—	—	.0374; .0377
Domain[a]	16	143	—	—	—	—	—	—	—	—	.006	16.164	—
Reading = Spelling = 0	—	—	2.440	—	—	—	—	.165	—	—	—	—	—
Reading vs. Spelling	—	—	0.854	—	—	—	—	.401	—	—	—	—	—
Reading	—	—	—	.048	.048	.027	1.800	.103	−.012	.108	—	—	.0481; .0497
Spelling	—	—	—	.120	.120	.074	1.620	.185	−.091	.331	—	—	.1196; .1203
Age	18	155	—	.001	.001	.004	0.154	.884	−.010	.011	.014	32.988	—
Sequence complexity	16	125	—	—	—	—	—	—	—	—	.018	37.434	—
FOC = SOC = 0	—	—	2.220	—	—	—	—	.177	—	—	—	—	—
FOC vs. SOC	—	—	0.692	—	—	—	—	.427	—	—	—	—	—
FOC	—	—	—	.028	.028	.029	0.972	.383	−.050	.105	—	—	.0277; .0282
SOC	—	—	—	.069	.069	.036	1.886	.098	−.016	.153	—	—	.0686; .0688
Session[b]	18	152	—	—	—	—	—	—	—	—	.008	21.693	—
T1 = T2 = 0	—	—	5.890	—	—	—	—	**.036**	—	—	—	—	—
T1 vs. T2	—	—	8.400	—	—	—	—	**.033**	—	—	—	—	—
T1	—	—	—	.023	.023	.024	0.947	.361	−.029	.075	—	—	—
T2	—	—	—	.164	.163	.045	3.646	**.025**	.035	.294	—	—	—
Number of trials	18	155	—	.000	.000	.000	1.118	.312	.000	.000	.011	28.483	4.43e-05; 4.23e-05
Type of SRTT	18	155	—	—	—	—	—	—	—	—	.016	35.618	—
Det = Prob = Alt = 0	—	—	0.967	—	—	—	—	.543	—	—	—	—	—
Det vs. Prob vs. Alt	—	—	0.291	—	—	—	—	.669	—	—	—	—	—
Deterministic	—	—	—	.046	.046	.023	2.041	.070	.005	.097	—	—	.0455; .0462
Probabilistic	—	—	—	.067	.067	.053	1.259	.308	−.117	.251	—	—	.0678; .0669
Alternating	—	—	—	.001	.001	.081	0.011	.993	−1.025	1.026	—	—	−.0004; .0012

Notes. F values are from Approximate Hotelling–Zhang with small sample correction omnibus tests of the effects of moderators with more than two levels; r = Pearson's R correlation; standard errors (SE) and t values for individual levels of a moderator; p values correspond to F or Fisher's z values; 95% CI corresponds to the Fisher's z.

Abbreviations: k, number of effect size estimates; s, number of studies.

[a]The effect of *subdomain* was only analyzed for spelling and reading.

[b]Given the small number of studies for the third session, this level was omitted from the analyses as the parameters could not be estimated.

$p = .414$). Again, there was no evidence supporting our hypothesis for a relationship between literacy and procedural memory in the typically developing group (H1), nor did the magnitude of the relationship between spelling and reading and procedural memory differed between the typically developing and dyslexic groups (H3). Yet, the disattenuated correlations (see Table 3) show a small relationship between spelling and procedural abilities for the typically developing group ($r = .10$), while for the DLD group, there was a small association between reading and procedural abilities ($R = .22$). Of particular interest, the spelling abilities of the dyslexic group moderately correlated with procedural learning ($R = .54$).

For all sampling and methodological moderators, the magnitude of the relationship was not significantly different between categorical levels nor from zero, except for *session*. The relationship between literacy and procedural learning was moderated by *session* ($F(2, 6.36) = 5.89$, $p = .036$), with a higher correlation between these variables for the second session (Fisher's $z = .164$, 95% CI [.035, .294] SE = .045, $t(3.67) = 3.65$, $p = .025$) than for the first (Fisher's $z = .023$, 95% CI [−.029, .075], SE = .024, $t(12.80) = .95$, $p = .361$). The difference in the magnitude of the effect size estimate between sessions 1 and 2 was statistically significant ($F(1, 5.13) = 8.40$, $p = .033$), with a higher effect for session 2 than session 1.

3.1.3.3. Model diagnostics: Sensitivity analyses show that the findings did not differ depending on the value of rho (see Tables 3 and 4). The study by West et al. (2021) was identified as a potential influential effect size, with an aggregated effect size higher than expected. Given that outlier detection was conducted on the aggregated data, no further actions were taken.

3.1.3.4. Publication bias: For publication bias, visual inspection of the funnel (shown in Fig. 3a) and contour (shown in Fig. 3b) plots of the aggregated effect sizes revealed no evidence of asymmetry. This is consistent with the nonsignificant rank test for plot asymmetry (Kendall's $\tau = -.026$, $p = .880$), thus suggesting low likelihood of publication bias.

PET and PEESE RVE models also showed no evidence of publication bias (PETrve: $b1 = .02$, $p = .961$; PEESErve: $b1 = .17$, $p = .840$).

4. Discussion

In the current study, we take a continuous approach to examining individual differences in literacy and language as predictors of procedural learning, in typical and literacy/language-disordered populations, as predicted by the procedural/declarative model. Our large-scale meta-analysis found limited evidence of support for the procedural/declarative model. Counter to this hypothesis (and our predictions) but in keeping with recent smaller-scale meta-analyses, the results revealed only a negligible relationship between procedural learning and language and literacy for the overall sample. Neither association remained significant when examined separately within each group (TD, DLD, and dyslexia). Turning to the separate subdomains of language and literacy, as expected, vocabulary did not significantly correlate with procedural learning for any of the groups, nor were there differences in the pooled effect

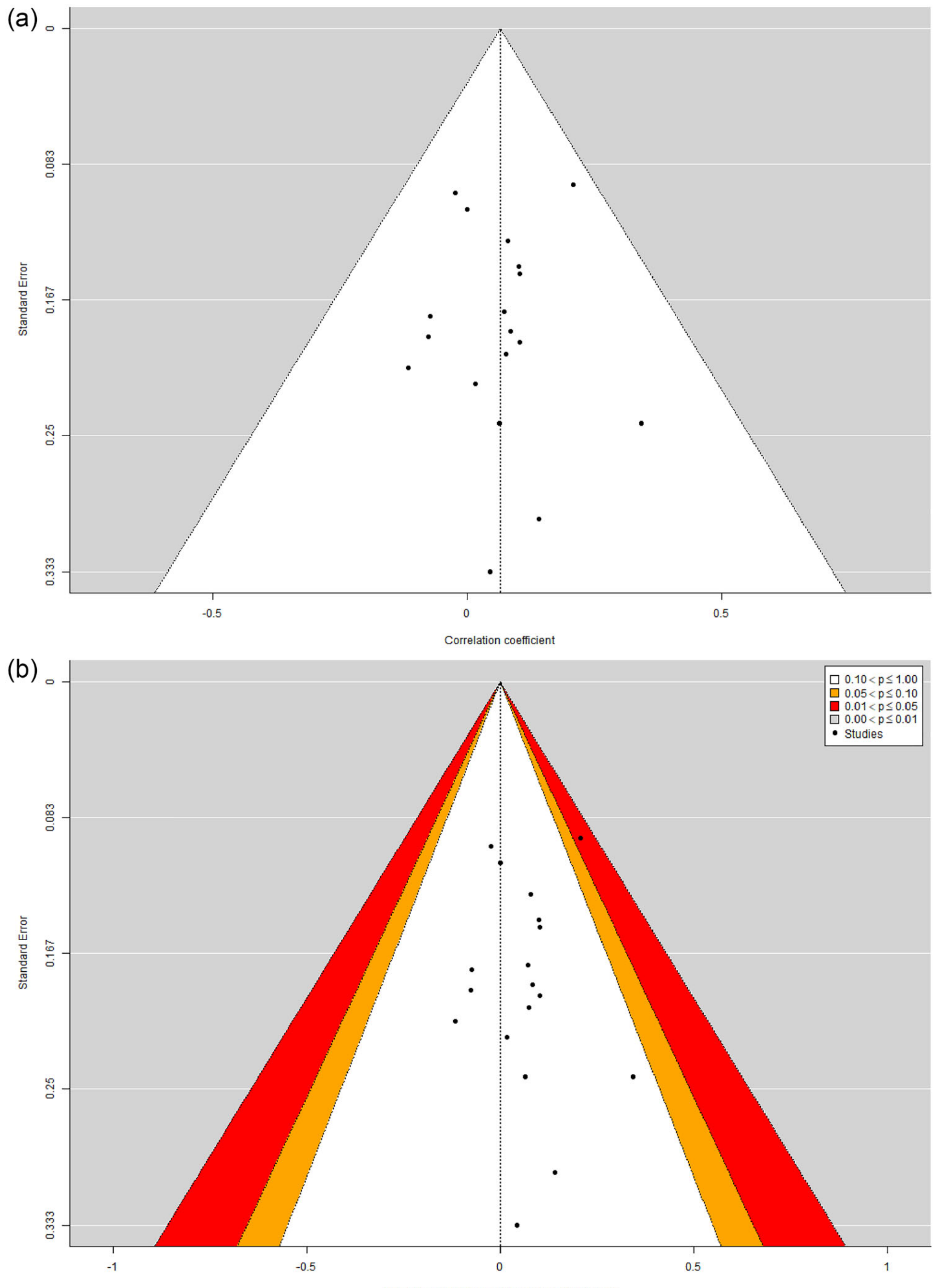

Fig. 3. Funnel plot showing study-level effect sizes plotted against standard error. An asymmetric distribution is taken as evidence of publication bias; (a) funnel plot (left panel) and (b) contour-enhanced funnel plot (right panel).

size between groups. However, there was also no evidence of a relationship between procedural learning and grammar, phonology, reading, and spelling in typically developing children, which is counter to our hypotheses and the predictions of the procedural/declarative model. Furthermore, the magnitude of the relationship between language and literacy and procedural learning did not differ between groups; specifically, the size of the effect did not differ between the typically developing and DLD groups for grammar and phonology, nor did it differ between typically developing and dyslexic groups for phonology, reading, and spelling. Together, these results provide minimal to no support for the procedural/declarative model or the procedural deficit hypothesis.

While the absence of correlations between language and literacy measures and procedural learning for the disordered groups may be taken as supportive evidence for the procedural deficit hypothesis (see Lum et al., 2012), this was observed alongside a similar pattern for the typically developing group. Thus, these results point to an overall lack of a continuous association between language, literacy, and procedural learning as previously observed by Lammertink et al. (2020) and West et al. (2021).

The findings from this meta-analysis replicate and extend the results of recent meta-analyses which found no association between grammar and procedural learning in children (Lammertink, Boersma, Wijnen, & Rispens, 2017) and between language and decoding measures and procedural learning on the SRTT (West et al., 2021). However, these findings are at odds with those obtained by Hamrick et al. (2018). The significant relationship between procedural learning and grammatical abilities observed by Hamrick et al. (2018) for first language learners may have been due to the small sample size, and inclusion of low-powered studies which have been found to often upwardly bias the overall estimate (Loken & Gelman, 2017; Turner, Bird, & Higgins, 2013). Importantly, subsequent larger published studies reported a nonsignificant relationship between procedural learning and grammatical abilities (e.g., Llompart & Dąbrowska, 2020; West et al., 2018). There was no evidence of a moderating role of age in the present meta-analysis, thus suggesting that the strength of the relationship between procedural learning and language/literacy was not influenced by the age of the participants counter to Hamrick et al. (2018) but in line with the finding from Lammertink et al. (2020). However, only children older than 5 years old were included in the present meta-analysis. Thus, we cannot rule out that procedural learning may be more tightly associated with language and literacy acquisition at earlier stages of development. The association between procedural learning and grammar abilities in adult second language learners observed by Hamrick et al. (2018) may capture this early stage of language acquisition, when linguistic rule-based knowledge is accumulated and integrated into more abstract and complex grammatical structures. Thus, it will be important for future work to take a broader age perspective. Another possible explanation for the discrepancy between the present results and Hamrick et al. (2018) is that the latter focused on studies that used deterministic SRTTs when examining the relationship between these variables in children, whereas the present review included both deterministic and probabilistic tasks; however, there were no differences in the effect sizes for these task variants here, making this explanation unlikely.

Furthermore, exploratory meta-regressions on the whole sample were conducted to assess the impact of methodological differences on the magnitude of the relationship between language and literacy and procedural learning. There was no evidence that sequence complexity, number of trials, and type of SRTT affected the association between language and literacy and procedural learning. Thus, even though the classic SRTT has been suggested to be a better index of the procedural memory function (Conway et al., 2019; Janacsek et al., 2020), there was no moderating effect of the type of SRTT on the relationship between procedural memory and language/literacy. However, the number of SRTT sessions was found to be a significant moderator of the relationship between procedural learning and literacy, such that the size of the effect was higher, even though still small, for procedural learning captured during a second session than for a first session. This is consistent with the suggestion by Conway et al. (2019), that correlations between literacy and procedural learning may emerge only for later sessions when procedural learning is more robust since knowledge and skill acquisition in this memory system tends to be gradual and require multiple exposures to the stimuli. That is, procedural learning that takes place after multiple training sessions may provide a more reliable predictor of individual differences than after a single session.

Importantly, the present findings suggest that while the SRTT is a robust experimental task which is able to capture group differences, it may not provide a reliable measure of individual differences (Hedge, Powell, & Sumner, 2018). In line with this, in their meta-analysis, West et al. (2021) found group differences between individuals with dyslexia and DLD and the typically developing group on procedural learning across measures (SRTT, Hebb learning task, artificial grammar and statistical learning tasks, weather prediction task). However, they found a nonsignificant relationship between continuous measures of language- and procedural learning. While a reliable measure of individual differences requires considerable interindividual variance allowing for the ranking of individuals, a sensitive measure of group differences does not. This interpretation concurs with recent evidence demonstrating the poor test-retest reliability of the procedural learning scores obtained with the SRTT (e.g., Kalra et al., 2019; Siegelman & Frost, 2015; Stark-Inbar et al., 2017; West et al., 2018; West et al., 2021). Given this, one must consider that the weak/absent correlations observed here may not necessarily refute the procedural/declarative model and the procedural deficit hypothesis, but rather, the SRTT may be insufficiently sensitive to individual differences to provide an adequate test of these hypotheses.

Test-retest reliability refers to the measure's consistency in ordering participants' performance at different time points (Kottner & Streiner, 2011). Measurement error and low variance between individuals have been suggested to decrease reliability (e.g., Fleiss, 1986; Hedge et al., 2018). Thus, the reliability of each measure will inform how much the raw correlations have been attenuated (e.g., Fleiss, 1986; Rouder et al., 2019; Spearman, 1904). The issue of attenuation has long been discussed (Spearman, 1904), yet, despite good progress, correlation recovery is still suboptimal as the methods available, while less biased than raw correlations, still produce highly variable estimates (Rouder et al., 2019). One such method was proposed by Spearman (1904) and it proposes that the disattenuation of a correlation between two measures can be accomplished by taking into account the reliability of each measure. We took this approach to estimate disattenuated correlations between language and literacy and procedural learning, but this did not change the pattern of results; correlations remained very low except for the correlations between procedural learning and phonology in DLD and spelling in dyslexia. Crucially, while disattenuated correlations may provide a better understanding of the true correlations, the Spearman (1904) method has been found to produce highly variable estimates and so these should be interpreted with caution. This is of special relevance given the pattern for low correlations for the typically developing group for which the procedural/declarative model makes clear predictions. Importantly, these results do not appear to be explained by publication bias. In sum, there appears to be some previous support for a procedural memory impairment in individuals with dyslexia and DLD in line with the procedural deficit hypothesis, as indicated by group-level studies. However, in the absence of evidence for a relationship between language and literacy and procedural memory as measured continuously-which may be a consequence of methodological limitations-it is still unclear whether and to what extent procedural memory underlies the development of language and literacy across the whole distribution of ability. While the SRTT has been shown to engage similar brain regions to motor skill learning (Keele, Ivry, Mayr, Hazeltine, & Heuer, 2003; Pascual-Leone, Wassermann, Grafman, & Hallett, 1996; Robertson et al., 2001; Torriero, Oliveri, Koch, Caltagirone, & Petrosini, 2004), and a more recent functional neuroanatomical meta-analysis by Janacsek et al. (2020) has clearly tied the SRTT to basal ganglia activation, little is known about whether the abilities required to perform the SRTT are predictive of how well an individual learns and performs real-world procedural learning tasks (Mathews, 1997) and indeed whether quantitative differences in procedural learning on the SRTT are meaningful. A related issue is the lack of correlation between different tasks purporting to measure procedural learning, even when task demands have been carefully matched (Erickson, Kaschak, Thiessen, & Berry, 2016). While this could again be explained by poor psychometric properties (Arnon, 2020; Siegelman & Frost, 2015; West et al., 2018; West et al., 2021), it is also possible that procedural learning is not a unified ability that can be similarly captured by all these tasks (Bogaerts et al., 2021). Thus, it may be that some measures of procedural learning are more relevant to the acquisition of language and literacy than others. This may explain the significant relationship between artificial grammar and statistical learning tasks and language-related abilities found by West et al. (2021). This challenges the view of procedural learning as a general capacity that underlies the acquisition of all probabilistic knowledge irrespective of modality and domain (Conway et al., 2019). As suggested by Bogaerts et al. (2021) and Siegelman, Bogaerts, Christiansen, and Frost (2017), future empirical work should focus on understanding the computations underlying procedural learning acquisition in each task so that these can be better mapped onto linguistic abilities. Thus, while this meta-analysis has focused on broad linguistic and literacy categories, future work should aim to map the relationship between procedural learning and specific linguistic processes and representations.

Finally, the absence of evidence for a relationship between procedural learning on the SRTT and language/literacy also raises the possibility that group differences on the SRTT may be unrelated to language and literacy skills. The SRTT is not a pure measure of procedural learning and has been shown to rely on attention and working memory (Arciuli, 2017; Sengottuvel & Rao, 2013; D. R. Shanks & St. John, 1994; West et al., 2021). Thus, in light of the evidence that individuals with dyslexia and DLD often have weaknesses in executive function (DLD: Marini, Piccolo, Taverna, Berginc, & Ozbič, 2020; dyslexia: Romani et al., 2011;

M. J. Snowling, Hulme, & Nation, 2020) and working memory (DLD: e.g., Baird et al., 2010; dyslexia: e.g., Fostick & Revah, 2018), group differences may actually reflect differences in other cognitive skills.

In conclusion, the present meta-analysis provides the most comprehensive overview to date of the relationship between procedural memory and language and literacy across children and adults with and without language and literacy disorders. The results provide limited evidence for a relationship between continuous measures of language and literacy and procedural learning as indexed by the SRTT, thus calling into question the validity of the procedural/declarative model and procedural deficit hypothesis as a framework for understanding language acquisition. However, this research is not without its limitations. As previously mentioned, even though the SRTT shows the most consistent neuroimaging evidence for the engagement of the basal ganglia (Janacsek et al., 2020), the poor psychometric properties of this task have likely downwardly biased the overall effect size. Thus, our efforts to correct for this attenuation should provide a better estimate of the true association between language/literacy and procedural memory (Wiernik & Dahlke, 2019). After disattenuation, there was a moderate association between spelling and procedural memory in the dyslexic group and between phonology and procedural memory in the DLD group, thus lending some weak support for the procedural deficit hypothesis. However, due to the lack of clear predictions for these groups, and the null effects for the control group, it is unclear whether these results support the involvement of procedural memory in language and literacy as proposed by the procedural/declarative model. The small sample sizes of the studies included in this meta-analysis and the age range of the participants may have also contributed to the null findings, particularly if developmental changes in this relationship are expected to occur. Finally, we have only partially assessed the predictions of the procedural/declarative model as the role of declarative memory in language and literacy has not been assessed. Thus, the inclusion of a declarative memory task would have been useful to better understand the dynamics between these long-term memory systems and their involvement in language and literacy acquisition and processing.

Future research is also needed to ascertain and improve the psychometric properties of the SRTT before this theoretical framework can be robustly tested. An important step will be for individual differences studies in this field to routinely adopt the practice of reporting test-retest reliability, allowing researchers to analyze the impact of reliability on their outcomes of interest (Parsons, Kruijt, & Fox, 2019). Additionally, more sophisticated models, such as meta-analytic structural equation modeling, may be better suited for assessing the relationship between language and literacy and procedural learning as latent variables, while taking measurement error into account. Such a meta-analysis should ideally include measures of procedural learning from multiple tasks that tap into different abilities across subdomains; as well as declarative memory and potential confounding variables, such as attention and working memory. This model would have the potential to shed light on the moderating and mediating effects of procedural learning on language and literacy in children and adults with and without neurodevelopmental disorders. Such future research endeavors will be important in advancing our understanding of procedural memory, and its putative role in language and literacy acquisition, with the potential for informing practice and intervention.

Acknowledgments

The authors would like to thank Dr. Gillian Clark and Dr. Kemény Ferenc for reviewing the preliminary list of relevant manuscripts and their extremely helpful suggestions. Also, we thank Professor Dorothy Bishop, Dr. Jisook Park, Dr. Sybren Spit, Professor Michael Ullman, Dr. Catherine Stoodley, Professor James Howard Jr, Dr. Joaquín M. M. Vaquero, Professor Stefano Vicari, Dr. Merel van Witteloostuijn, Dr. Eliane Roulet-Perez, Dr. Prema Rao, Professor Jarrad Lum, Dr. Joanna Chen Lee, Dr. Julie Hsu, Dr. Deny Menghini, Dr. Floriana Costanzo, Dr. Ferenc Kemény, Professor Eli Vakil, Dr. Steve Kelly, Dr. Martina Hedenius, Dr. Yafit Gabay, Dr. Nicolas Stefaniak, Dr. Dorthe Bleses, Dr. Evan Kidd, Dr. Frances Sayako Earle, and Dr. Imme Lammertink for providing additional data and information that enabled us to conduct the meta-analysis.

Conflict of interest

There is no relevant conflict of interest.

Open Research Badges

This article has earned Open Data and Open Materials badges. Data and materials are available at https://osf.io/ev2xw/.

Note

1 Other subdomains were coded (e.g., reading comprehension), yet due to their small sample size, separate analyses for these subdomains were not conducted. Instead, they have only been included in the analyses for the overall effect of language or literacy.

References

References marked with an asterisk (*) were included in the meta-analysis.

Arciuli, J. (2017). The multi-component nature of statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160058. https://doi.org/10.1098/rstb.2016.0058

Arnon, I. (2019). Statistical Learning, Implicit Learning, and First Language Acquisition: A Critical Evaluation of Two Developmental Predictions. *Topics in Cognitive Science*. https://doi.org/10.1111/tops.12428

Arnon, I. (2020). Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behavior Research Methods*, *52*, (1), 68–81. https://doi.org/10.3758/s13428-019-01205-5

Baird, G., Dworzynski, K., Slonims, V., & Simonoff, E. (2010). Memory impairment in children with language impairment. *Developmental Medicine and Child Neurology*, *52*(6), 535–540. https://doi.org/10.1111/j.1469-8749.2009.03494.x

Batterink, L. J., Paller, K. A., & Reber, P. J. (2019). Understanding the neural bases of implicit and statistical learning. *Topics in Cognitive Science*, *11*, 482–503. https://doi.org/10.1111/tops.12420

Barker, L. (2012). Defining the parameters of incidental learning on a serial reaction time (SRT) task: Do conscious rules apply? *Brain Sciences*, *2*(4), 769–789. https://doi.org/10.3390/brainsci2040769

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*(4), 1088. https://doi.org/10.2307/2533446

Bogaerts, L., Siegelman, N., Ben-Porat, T., & Frost, R. (2018). Is the Hebb repetition task a reliable measure of individual differences in sequence learning? *Quarterly Journal of Experimental Psychology*, *71*(4), 892–905. https://doi.org/10.1080/17470218.2017.1307432

Bogaerts, L., Siegelman, N., & Frost, R. (2021). Statistical learning and language impairments: Toward more precise theoretical accounts. *Perspectives on Psychological Science*, *16*(2), 319–337. https://doi.org/10.1177/1745691620953082

Borestein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *Handbook of research synthesis and meta-analysis* (pp. 221–236). Russell Sage Foundation.

Brookman, A., McDonald, S., McDonald, D., & Bishop, D. V. M. (2013). Fine motor deficits in reading disability and language impairment: Same or different? *PeerJ*, *1*, e217. https://doi.org/10.7717/peerj.217

Buchholz, J., & McKone, E. (2004). Adults with dyslexia show deficits on spatial frequency doubling and visual attention tasks. *Dyslexia*, *10*(1), 24–43. https://doi.org/10.1002/dys.263

Christiansen, M. H., Conway, C. M., & Onnis, L. (2012). Similar neural correlates for language and sequential learning: Evidence from event-related brain potentials. *Language and Cognitive Processes*, *27*(2), 231–256. https://doi.org/10.1080/01690965.2011.606666

*Clark, G. M., & Lum, J. A. G. (2017a). Procedural memory and speed of grammatical processing: Comparison between typically developing children and language impaired children. *Research in Developmental Disabilities*, *71*, 237–247. https://doi.org/10.1016/j.ridd.2017.10.015

* Clark, G. M., & Lum, J. A. G. (2017b). First-order and higher order sequence learning in specific language impairment. *Neuropsychology*, *31*(2), 149–159. https://doi.org/10.1037/neu0000316

Conway, C. M., Arciuli, J., Lum, J. A. G., & Ullman, M. T. (2019). Seeing problems that may not exist: A reply to West et al.'s (2018) questioning of the procedural deficit hypothesis. *Development Science*, *22*(4), e12814.

Conway, C. M., & Pisoni, D. B. (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Annals of the New York Academy of Sciences*, *1145*(1), 113–131. https://doi.org/10.1196/annals.1416.009

Del Re, A., & Hoyt, W. T. (2018). *MAc: Meta-analysis with correlations*. Retrieved from https://CRAN.R-project.org/package=MAc

Delage, H., & Frauenfelder, U. H. (2020). Relationship between working memory and complex syntax in children with developmental language disorder. *Journal of Child Language*, *47*(3), 600–632. https://doi.org/10.1017/S0305000919000722

*Deroost, N., Zeischka, P., Coomans, D., Bouazza, S., Depessemier, P., & Soetens, E. (2010). Intact first- and second-order implicit sequence learning in secondary-school-aged children with developmental dyslexia. *Journal of Clinical and Experimental Neuropsychology*, *32*(6), 561–572. https://doi.org/10.1080/13803390903313556

*Desmottes, L., Meulemans, T., & Maillart, C. (2016). Later learning stages in procedural memory are impaired in children with specific language impairment. *Research in Developmental Disabilities*, *48*, 53–68. https://doi.org/10.1016/j.ridd.2015.10.010

*Desmottes, L., Maillart, C., & Meulemans, T. (2017a). Memory consolidation in children with specific language impairment: Delayed gains and susceptibility to interference in implicit sequence learning. *Journal of Clinical and Experimental Neuropsychology*, *39*(3), 265–285. https://doi.org/10.1080/13803395.2016.1223279

*Desmottes, L., Meulemans, T., Patinec, M. A., & Maillart, C. (2017b). Distributed training enhances implicit sequence acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *60*(9), 2636–2647. https://doi.org/10.1044/2017_JSLHR-L-16-0146

Earle, F. S., & Ullman, M. T. (2021). Deficits of Learning in Procedural Memory and Consolidation in Declarative Memory in Adults With Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, *64*(February), 1–11. https://doi.org/10.1044/2020_jslhr-20-00292

Erickson, L. C., Kaschak, M. P., Thiessen, E. D., & Berry, C. A. S. (2016). Individual differences in statistical learning: Conceptual and measurement issues. *Collabra*, *2*(1), 14. https://doi.org/10.1525/collabra.41

Fawcett, A. J., Nicolson, R. I., & Dean, P. (1996). Impaired performance of children with dyslexia on a range of cerebellar tasks. *Annals of Dyslexia*, *46*(1), 259–283. https://doi.org/10.1007/BF02648179

Fisher, Z., & Tipton, E. (2015). robumeta: An R-package for robust variance estimation in meta-analysis. *ArXiv:1503.02220 [Stat]*. Retrieved from https://arxiv.org/abs/1503.02220

Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. John Wiley & Sons.

Fostick, L., & Revah, H. (2018). Dyslexia as a multi-deficit disorder: Working memory and auditory temporal processing. *Acta Psychologica*, *183*, 19–28. https://doi.org/10.1016/j.actpsy.2017.12.010

*Gabay, Y., Schiff, R., & Vakil, E. (2012). Dissociation between the procedural learning of letter names and motor sequences in developmental dyslexia. *Neuropsychologia*, *50*(10), 2435–2441. https://doi.org/10.1016/j.neuropsychologia.2012.06.014

*Gabriel, A., Maillart, C., Guillaume, M., Stefaniak, N., & Meulemans, T. (2011). Exploration of serial structure procedural learning in children with language impairment. *Journal of the International Neuropsychological Society*, *17*(2), 336–343. https://doi.org/10.1017/S1355617710001724

*Gabriel, A., Maillart, C., Stefaniak, N., Lejeune, C., Desmottes, L., & Meulemans, T. (2013). Procedural learning in specific language impairment: Effects of sequence complexity. *Journal of the International Neuropsychological Society*, *19*(3), 264–271. https://doi.org/10.1017/S1355617712001270

*Gabriel, A., Meulemans, T., Parisse, C., & Maillart, C. (2015). Procedural learning across modalities in French-speaking children with specific language impairment. *Applied Psycholinguistics*, *36*(3), 747–769. https://doi.org/10.1017/S0142716413000490

*Gabriel, A., Stefaniak, N., Maillart, C., Schmitz, X., & Meulemans, T. (2012). Procedural visual learning in children with specific language impairment. *American Journal of Speech-Language Pathology*, *21*(4), 329–341. https://doi.org/10.1044/1058-0360(2012/11-0044)

Galbraith, R. F. (1988). A note on graphical presentation of estimated odds ratios from several clinical trials. *Statistics in medicine*, *7*(8), 889–894.

Hamrick, P., Lum, J. A. G., & Ullman, M. T. (2018). Child first language and adult second language are both tied to general-purpose learning systems. *Proceedings of the National Academy of Sciences*, *115*(7), 1487–1492. https://doi.org/10.1073/pnas.1713975115

*Hedenius, M., Lum, J. A. G., & Bölte, S. (2021). Alterations of procedural memory consolidation in children with developmental dyslexia. *Neuropsychology*, *35*(2), 185–196. https://doi.org/10.1037/neu0000708

*Hedenius, M., Persson, J., Alm, P. A., Ullman, M. T., Howard, J. H., Howard, D. V., & Jennische, M. (2013). Impaired implicit sequence learning in children with developmental dyslexia. *Research in Developmental Disabilities*, *34*(11), 3924–3935. https://doi.org/10.1016/j.ridd.2013.08.014

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39–65. https://doi.org/10.1002/jrsm.5

*Henderson, L. M., & Warmington, M. (2017). A sequence learning impairment in dyslexia? It depends on the task. *Research in Developmental Disabilities*, *60*, 198–210. https://doi.org/10.1016/j.ridd.2016.11.002

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Hill, E. L. (2001). Non-specific nature of specific language impairment: A review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders*, *36*(2), 149–171. https://doi.org/10.1080/13682820010019874

*Hsu, H. J., & Bishop, D. V. M. (2014). Sequence-specific procedural learning deficits in children with specific language impairment. *Developmental Science*, *17*(3), 352–365. https://doi.org/10.1111/desc.12125

Jackson, E., Leitão, S., Claessen, M., & Boyes, M. (2020). Working, declarative, and procedural memory in children with developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *63*(12), 4162–4178. https://doi.org/10.1044/2020_JSLHR-20-00135

Janacsek, K., Shattuck, K. F., Tagarelli, K. M., Lum, J. A. G., Turkeltaub, P. E., & Ullman, M. T. (2020). Sequence learning in the human brain: A functional neuroanatomical meta-analysis of serial reaction time studies. *NeuroImage*, *207*, Article 116387. https://doi.org/10.1016/j.neuroimage.2019.116387

Kalra, P. B., Gabrieli, J. D. E., & Finn, A. S. (2019). Evidence of stable individual differences in implicit learning. *Cognition*, *190*, 199–211. https://doi.org/10.1016/j.cognition.2019.05.007

Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review*, *110*(2), 316–339. https://doi.org/10.1037/0033-295X.110.2.316

*Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*(1), 171–184. https://doi.org/10.1037/a0025405

*Kidd, E., & Kirjavainen, M. (2011). Investigating the contribution of procedural and declarative memory to the acquisition of past tense morphology: Evidence from Finnish. *Language and Cognitive Processes*, *26*(4–6), 794–829. https://doi.org/10.1080/01690965.2010.493735

Kirkham, J. J., Dwan, K. M., Altman, D. G., Gamble, C., Dodd, S., Smyth, R., & Williamson, P. R. (2010). The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ*, *340*, c365. https://doi.org/10.1136/bmj.c365

Krishnan, S., Watkins, K. E., & Bishop, D. V. M. (2016). Neurobiological basis of language learning difficulties. *Trends in Cognitive Sciences*, *20*(9), 701–714. https://doi.org/10.1016/j.tics.2016.06.012

Kottner, J., & Streiner, D. L. (2011). The difference between reliability and agreement. *Journal of Clinical Epidemiology*, *64*(6), 701–702. https://doi.org/10.1016/j.jclinepi.2010.12.001

*Kuppuraj, S., Rao, P., & Bishop, D. V. (2016). Declarative capacity does not trade-off with procedural capacity in children with specific language impairment. *Autism & Developmental Language Impairments*, *1*, 1–17. https://doi.org/10.1177/2396941516674416

*Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2017). Statistical learning in specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *60*(12), 3474–3486. https://doi.org/10.1044/2017_JSLHR-L-16-0439

Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2020). Statistical learning in the visuomotor domain and its relation to grammatical proficiency in children with and without developmental language disorder: A conceptual replication and meta-analysis. *Language Learning and Development*, *16*(4), 426–450. https://doi.org/10.1080/15475441.2020.1820340

*Lee, J. C., & Tomblin, J. B. (2015). Procedural learning and individual differences in language. *Language Learning and Development*, *11*(3), 215–236. https://doi.org/10.1080/15475441.2014.904168

Light, R. J., Richard, J., Light, R., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research*. Harvard University Press.

*Llompart, M., & Dąbrowska, E. (2020). Explicit but not implicit memory predicts ultimate attainment in the native language. *Frontiers in Psychology*, *11*, 1–14. https://doi.org/10.3389/fpsyg.2020.569586

Loken, E., & Gelman, A. (2017). Measurement error and the replication crisis. *Science*, *355*(6325), 584–585. https://doi.org/10.1126/science.aal3618

*Lukács, Á., & Kemény, F. (2014). Domain-general sequence learning deficit in specific language impairment. *Neuropsychology*, *28*(3), 472–483. https://doi.org/10.1037/neu0000052

Lum, J. A. G., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex*, *51*(100), 1–10. https://doi.org/10.1016/j.cortex.2013.10.011

*Lum, J. A. G., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex*, *48*(9), 1138–1154. https://doi.org/10.1016/j.cortex.2011.06.001

*Lum, J. A. G., & Kidd, E. (2012). An examination of the associations among multiple memory systems, past tense, and vocabulary in typically developing 5-year-old children. *Journal of Speech, Language, and Hearing Research*, *55*(4), 989–1006. https://doi.org/10.1044/1092-4388(2011/10-0137)

Lum, J. A. G., Ullman, M. T., & Conti-Ramsden, G. (2013). Procedural learning is impaired in dyslexia: Evidence from a meta-analysis of serial reaction time studies. *Research in Developmental Disabilities*, *34*(10), 3460–3476. https://doi.org/10.1016/j.ridd.2013.07.017

Marini, A., Piccolo, B., Taverna, L., Berginc, M., & Ozbič, M. (2020). The complex relation between executive functions and language in preschoolers with developmental language disorders. *International Journal of Environmental Research and Public Health*, *17*(5), 1772. https://doi.org/10.3390/ijerph17051772

Mathews, R. C. (1997). Is research painting a biased picture of implicit learning? The dangers of methodological purity in scientific debate. *Psychonomic Bulletin & Review*, *4*(1), 38–42. https://doi.org/10.3758/BF03210771

*Mayor-Dubois, C., Zesiger, P., Van der Linden, M., & Roulet-Perez, E. (2014). Nondeclarative learning in children with specific language impairment: Predicting regularities in the visuomotor, phonological, and cognitive domains. *Child Neuropsychology*, *20*(1), 14–22. https://doi.org/10.1080/09297049.2012.734293

*Menghini, D., Finzi, A., Benassi, M., Bolzani, R., Facoetti, A., Giovagnoli, S., Ruffino, M., & Vicari, S. (2010). Different underlying neurocognitive deficits in developmental dyslexia: A comparative study. *Neuropsychologia*, *48*(4), 863–872.

*Mimeau, C., Coleman, M., & Donlan, C. (2016). The role of procedural memory in grammar and numeracy skills. *Journal of Cognitive Psychology*, *28*(8), 899–908.

Nicolson, R. I., & Fawcett, A. J. (2011). Dyslexia, dysgraphia, procedural learning and the cerebellum. *Cortex*, *47*(1), 117–127. https://doi.org/10.1016/j.cortex.2009.08.016

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1–32. https://doi.org/10.1016/0010-0285(87)90002-8

Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, *3*(1), 1–18. https://doi.org/10.1016/0022-2496(66)90002-2

*Oliveira, C. M., Hayiou-Thomas, M. E., & Henderson, L. M. (2022, May 10). Reliability of the Serial Reaction Time task: If at first you don't succeed, try try try again. PsyArxiv. https://doi.org/10.31234/osf.io/hqmy7

Oliveira, C. M., Hayiou-Thomas, M. E., & Henderson, L. M. (2023, April 25). The reliability of the serial reaction time task: Meta-analysis of test–retest correlation. PsyArxiv. https://doi.org/10.31234/osf.io/bjwqk

Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience*, *25*(1), 563–593. https://doi.org/10.1146/annurev.neuro.25.112701.142937

Park, J., Miller, C. A., Rosenbaum, D. A., Sanjeevan, T., van Hell, J. G., Weiss, D. J., & Mainela-Arnold, E. (2018). Bilingualism and procedural learning in typically developing children and children with language impairment. *Journal of Speech, Language, and Hearing Research*, *61*(3), 634–644. https://doi.org/10.1044/2017_JSLHR-L-16-0409

Parsons, S., Kruijt, A.-W., & Fox, E. (2019). Psychological science needs a standard practice of reporting the reliability of cognitive-behavioral measurements. *Advances in Methods and Practices in Psychological Science*, *2*(4), 378–395. https://doi.org/10.1177/2515245919879695

Pascual-Leone, A., Wassermann, E. M., Grafman, J., & Hallett, M. (1996). The role of the dorsolateral prefrontal cortex in implicit procedural learning. *Experimental Brain Research*, *107*(3), 479–485. https://doi.org/10.1007/BF00230427

Pustejovsky, J. (2021). clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections. Retrieved from https://CRAN.R-project.org/package=clubSandwich

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with Robust Variance Estimation: Expanding the Range of Working Models. *Prevention Science*, *23*(3), 425–438. https://doi.org/10.1007/s11121-021-01246-3

Robertson, E. M., Tormos, J. M., & Maeda, F. (2001). The role of the dorsolateral prefrontal cortex during sequence learning is specific for spatial information. *Cerebral Cortex*, 11(7), 628–635.

Romani, C., Tsouknida, E., di Betta, A. M., & Olson, A. (2011). Reduced attentional capacity, but normal processing speed and shifting of attention in developmental dyslexia: Evidence from a serial task. *Cortex*, *47*(6), 715–733. https://doi.org/10.1016/j.cortex.2010.05.008

Rouder, J. N., Kumar, A., & Haaf, J. M. (2019, March 25). Why most studies of individual differences with inhibition tasks are bound to fail. PsyArxiv. https://doi.org/10.31234/osf.io/3cjr5

Saffran, J. R. (2020). Statistical language learning in infancy. *Child Development Perspectives*, *14*(1), 49–54.

*Schmalz, X., Moll, K., Mulatti, C., & Schulte-Körne, G. (2019). Is statistical learning ability related to reading ability, and if so, why? *Scientific Studies of Reading*, *23*(1), 64–76. https://doi.org/10.1080/10888438.2018.1482304

Schwarb, H., & Schumacher, E. (2012). Generalized lessons about sequence learning from the study of the serial reaction time task. *Advances in Cognitive Psychology*, *8*(2), 165–178. https://doi.org/10.5709/acp-0113-1

Sengottuvel, K., & Rao, P. K. S. (2013). An adapted serial reaction time task for sequence learning measurements. *Psychological Studies*, *58*(3), 276–284. https://doi.org/10.1007/s12646-013-0204-z

Shanks, D. R., & St John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*(3), 367–395. https://doi.org/10.1017/S0140525X00035032

Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences* B, *372*, 20160059. https://doi.org/10.1098/rstb.2016.0059

* Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. https://doi.org/10.1016/j.jml.2015.02.001

Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word–referent learning. *Trends in Cognitive Sciences*, *18*(5), 251–258.

Snowling, M. (2000). *Dyslexia* (2nd ed). Blackwell.

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

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, *15*(2), 201–292.

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60–78. https://doi.org/10.1002/jrsm.1095

Stark-Inbar, A., Raza, M., Taylor, J. A., & Ivry, R. B. (2017). Individual differences in implicit motor learning: Task specificity in sensorimotor adaptation and sequence learning. *Journal of Neurophysiology*, *117*(1), 412–428. https://doi.org/10.1152/jn.01141.2015

*Spit, S., & Rispens, J. (2019). On the relation between procedural learning and syntactic proficiency in gifted children. *Journal of Psycholinguistic Research*, *48*(2), 417–429. https://doi.org/10.1007/s10936-018-9611-6

*Stoodley, C. J., Harrison, E. P. D., & Stein, J. F. (2006). Implicit motor learning deficits in dyslexic adults. *Neuropsychologia*, *44*(5), 795–798. https://doi.org/10.1016/j.neuropsychologia.2005.07.009

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life-Course Criminology*, *2*(1), 85–112. https://doi.org/10.1007/s40865-016-0026-5

Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics*, *40*(6), 604–634. https://doi.org/10.3102/1076998615606099

Torriero, S., Oliveri, M., Koch, G., Caltagirone, C., & Petrosini, L. (2004). Interference of left and right cerebellar rTMS with procedural learning. *Journal of Cognitive Neuroscience*, *16*(9), 1605–1611. https://doi.org/10.1162/0898929042568488

Turner, R. M., Bird, S. M., & Higgins, J. P. T. (2013). The impact of study size on meta-analyses: Examination of underpowered studies in Cochrane reviews. *PLoS ONE*, *8*(3), e59202. https://doi.org/10.1371/journal.pone.0059202

Ullman, M. T. (2001a). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, *30*(1), 37–69. https://doi.org/10.1023/A:1005204207369

Ullman, M. T. (2001b). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience*, *2*(10), 717–726. https://doi.org/10.1038/35094573

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1–2), 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Ullman, M. T. (2015). The declarative/procedural model: A neurobiologically motivated theory of first and second language. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd edition, pp. 135–158). Routledge.

Ullman, M. T. (2016a). The declarative/procedural model. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 953–968). Elsevier. https://doi.org/10.1016/B978-0-12-407794-2.00076-6

Ullman, M. T. (2016b). Chapter 76—The declarative/procedural model: A neurobiological model of language learning, knowledge, and use. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 953–968). Academic Press. https://doi.org/10.1016/B978-0-12-407794-2.00076-6

Ullman, M. T., Earle, F. S., Walenski, M., & Janacsek, K. (2020). The neurocognition of developmental disorders of language. *Annual Review of Psychology*, *71*(1), 389–417. https://doi.org/10.1146/annurev-psych-122216-011555

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, *41*(3), 399–433. https://doi.org/10.1016/S0010-9452(08)70276-4

Ullman, M. T., & Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews*, *51*, 205–222. https://doi.org/10.1016/j.neubiorev.2015.01.008

* Vakil, E., Lowe, M., & Goldfus, C. (2015). Performance of children with developmental dyslexia on two skill learning tasks—Serial Reaction Time and Tower of Hanoi Puzzle: A test of the specific procedural learning difficulties theory. *Journal of Learning Disabilities*, *48*(5), 471–481. https://doi.org/10.1177/0022219413508981

van der Lely, H. K. J. (2005). Domain-specific cognitive systems: Insight from Grammatical-SLI. *Trends in Cognitive Sciences*, *9*(2), 53–59. https://doi.org/10.1016/j.tics.2004.12.002

van der Lely, H. K. J., & Pinker, S. (2014). The biological basis of language: Insight from developmental grammatical impairments. *Trends in Cognitive Sciences*, *18*(11), 586–595. https://doi.org/10.1016/j.tics.2014.07.001

* van Witteloostuijn, M., Boersma, P., Wijnen, F., & Rispens, J. (2019). Statistical learning abilities of children with dyslexia across three experimental paradigms. *PLoS ONE*, *14*(8), Article e0220041. https://doi.org/10.1371/journal.pone.0220041

Vicari, S., Marotta, L., Menghini, D., Molinari, M., & Petrosini, L. (2003). Implicit learning deficit in children with developmental dyslexia. *Neuropsychologia*, *41*(1), 108–114. https://doi.org/10.1016/S0028-3932(02)00082-9

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

von Hippel, P. T. (2015). The heterogeneity statistic I^2 can be biased in small meta-analyses. *BMC Medical Research Methodology*, *15*, (1), 35. https://doi.org/10.1186/s12874-015-0024-z

*West, G., Clayton, F. J., Shanks, D. R., & Hulme, C. (2019). Procedural and declarative learning in dyslexia. *Dyslexia*, *25*(3), 246–255. https://doi.org/10.1002/dys.1615

West, G., Melby-Lervåg, M., & Hulme, C. (2021). Is a procedural learning deficit a causal risk factor for developmental language disorder or dyslexia? A meta-analytic review. *Developmental Psychology*, *57*(5), 749–770. https://doi.org/10.1037/dev0001172

*West, G., Shanks, D. R., & Hulme, C. (2021). Sustained attention, not procedural learning, is a predictor of reading, language and arithmetic skills in children. *Scientific Studies of Reading*, *25*(1), 47–63. https://doi.org/10.1080/10888438.2020.1750618

*West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2018). The procedural learning deficit hypothesis of language learning disorders: We see some problems. *Developmental Science*, *21*(2), e12552. https://doi.org/10.1111/desc.12552

Wiernik, B. M., & Dahlke, J. A. (2019). Obtaining unbiased results in meta-analysis: The importance of correcting for statistical artefacts [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/9mpbn

Williams, J. N. (2020). The neuroscience of implicit learning. *Language Learning*, *70*(S2), 255–307. https://doi.org/10.1111/lang.12405