Title: Lexical Precision Moderates Stroop Interference in Dyslexia and Stuttering

Mahmoud Elsherif[1,2], James H. Smith-Spark[3], Linda Wheeldon[4], Steven Frisson[1]

[1] School of Psychology, University of Birmingham, UK.

[2] School of Psychology and Vision Sciences, University of Leicester, UK.

mahmoud.medhat.elsherif@gmail.com ORCID: https://orcid.org/0000-0002-0540-3998

[3] School of Allied Health and Life Sciences, London South Bank University, London, UK

[4] Department of Foreign Languages and Translation, University of Agder, Kristiansand, Norway

Corresponding Author: Mahmoud Elsherif, School of Psychology and Vision Sciences, University of Leicester, UK. mahmoud.medhat.elsherif@gmail.com

ORCID iD: Mahmoud Elsherif: https://orcid.org/0000-0002-0540-3998

Abstract
Purpose: The Lexical Quality Hypothesis (LQH) posits that differences in the precision of orthographic, semantic, and phonological word representations affect language processing. While LQH components are typically investigated in reading, their role in resolving cognitive conflict remains unknown. This study applies the construct of lexical precision to the traditional naming Stroop task, a gold-standard to quantify cognitive conflict resolution specifically investigating its influence in dyslexia and stuttering.
Method: Stroop interference was measured in 84 neurotypical readers, 50 dyslexic adults and 30 adults who stutter (AWS). All participants completed a naming Stroop task, together with a suite of individual difference measures that were analysed using a Principal Component Analysis, which established measures of phonological precision and lexical precision.
Results: The data showed that, on average, dyslexic adults and AWS showed significantly larger Stroop interference effects compared with neurotypical adults. Crucially, the magnitude of the Stroop interference was moderated by lexical precision. Dyslexic adults and AWS who scored higher on the lexical precision component showed Stroop interference levels similar to the neurotypical adults, while those with lower precision demonstrated a significantly larger Stroop interference. However, the level of Stroop interference did not differ between AWS and dyslexic adults.
Conclusions: The results indicate that an imprecise lexical representation underpins the difficulties in cognitive control mechanisms that are independent of the specific neurodivergent groups.

Keywords: Dyslexia; Lexical Precision; Lexical Quality Hypothesis; Stroop Effect; Stuttering

Background

Word recognition is a paradoxical process. On the one hand, reading is, at face value, an automatic, effortless process, allowing cognitive resources to be spent on higher-order processes such as reading comprehension (LaBerge & Samuels, 1974). On the other hand, reading involves controlled processes, including the ability to suppress multiple words that compete with the target for activation in order to select the target word correctly. For instance, the error made in saying "dog" or "braked" instead of "cat" or "broke" can be attributed to competition between semantically or form-related words (Dell, 1986), respectively. Such competition resolution itself is an aspect of executive function (e.g., Sommers & Danielson, 1999) and can be measured using the Stroop interference task (Stroop, 1935). In this task, the font colour in which a word is presented needs to be named, while the reading of the word name itself needs to be suppressed. The level of effectiveness in suppressing competitors can differ greatly between people, with, for instance, some adults experiencing challenges when attempting to suppress competitors (e.g., dyslexic adults: Proulx & Elmasry, 2015; adults who stutter: Treleaven & Coalson, 2020). Executive function skills, especially the ability to suppress competitors for languages, are associated with psychosocial development, mental health, and overall quality of life across the lifespan (e.g., Best et al., 2011; Brown & Landgraf, 2010; Garon et al., 2008; Gathercole et al., 2004; Ofoe et al., 2018, 2025). In the current study, we investigated the relationship between language ability and executive function, using the Stroop interference task, in dyslexic individuals, people who stutter (PWS) and typical readers.

While definitions vary, executive function can be broadly defined as a set of higher-order cognitive control mechanisms that are responsible for goal-directed behaviour (Diamond, 2013; Miyake et al., 2000). From a theoretical perspective, updating, set shifting, and inhibition have been identified as “core” executive functions (e.g., Diamond, 2013; Miyake et al., 2000), upon which more sophisticated executive functions can be built (such as planning;

Miyake & Friedman, 2012). Within this framework, updating reflects the ability to update the contents of working memory in the light of newly available information, set shifting relates to cognitive flexibility in switching between operations or perspectives, and inhibition describes the ability to suppress habitual or pre-potent responses in favour of responses that are more appropriate to the current context or task demands. Of direct relevance to the current paper, deficits have been found in these core executive functions in neurodivergent populations, including dyslexic adults[1] (e.g., Henry et al., 2017, Smith-Spark et al., 2016) and children who stutter (CWS; e.g., Kakuta & Kawasaki, 2025), the foci of the current study.

Developmental dyslexia (henceforth, dyslexia) is characterised by a neurodivergence that manifests itself as a form of reading and spelling impairment (e.g., Vellutino et al., 2004). It is the most commonly occurring neurodevelopmental condition, with estimated prevalence rates estimated at 7 to 10% of the population (e.g., Peterson & Pennington, 2012). Indeed, dyslexia constitutes approximately 80% of all neurodevelopmental diagnoses (Handler et al., 2011). While there are many different theories regarding the causes of dyslexia (e.g., see O'Brien & Yeatman, 2021; Ramus et al., 2003; Stein, 2023), one predominant argument is that dyslexic individuals have a phonological processing impairment, with individuals struggling to connect sounds to letters, leading to the core phonological processing deficit theory of dyslexia (for reviews, see Castles & Friedmann, 2014; Vellutino et al., 2004). Beyond these core impairments in phonological processing, dyslexic children have been found to have broader cognitive deficits, such as difficulties with executive function (for reviews, see Booth et al., 2010; Lonergan et al., 2019) and memory (e.g., Jeffries & Everatt, 2004; Khan, 2014; Menghini et al., 2011). These broader cognitive difficulties continue in adulthood, with dyslexic adults showing impairments under laboratory conditions in executive function (e.g., Brosnan et al., 2002; Smith-Spark et al., 2016) and across a range of memory systems (e.g., Provazza et al., 2019; Smith-Spark & Fisk, 2007; Smith-Spark,

[1] Although there is a debate concerning person-first and identity-first language within academia and the neurodivergent community, recent research has shown that, on average, dyslexic adults prefer identity-first language over person-first language and adults who stutter prefer person-first language over identity-first language (Pearson et al., 2025).

Ziecik et al., 2017). In relation to conflict resolution, dyslexic-related executive function impairments have also been observed in response to non-verbal tasks where stronger interference effects have been found than in neurotypical individuals (e.g., Bednarek et al., 2004; Facoetti & Turatto, 2000; Mahe et al., 2014). Despite such wider cognitive difficulties being well documented empirically, they remain challenging for current theories to accommodate (see Smith-Spark & Gordon, 2022). This is particularly the case for explanations that are centred on core phonological processing difficulties (e.g., Vellutino, 1979). With particular regard to dyslexia and executive function, Smith-Spark et al. (2017) have, moreover, highlighted the challenges involved in disentangling executive problems from phonological processing difficulties when interpreting performance on executive fluency tasks since many tasks designed to assess executive function have a phonological processing component.

Developmental stuttering (henceforth, stuttering) is defined as an impairment in the ability to formulate motor plans for speech production. Stuttering occurs in about 8.5% of children and 0.81% of adults (Yairi & Ambrose, 2013). While the exact mechanisms are unknown, several models describe the multifactorial and dynamic nature (e.g., Smith & Weber, 2017), several models posit that the traits of stuttering and the level of support are multifactorial and dynamic, being associated with varying levels of cognitive, language-specific, and emotional factors (Bowers et al., 2018). Executive function problems have been observed in CWS (e.g., Anderson & Ofoe, 2019; Eggers et al., 2012, 2013; Eichorn et al., 2018) and adults who stutter (AWS; e.g., Lescht et al., 2024). These include challenges with inhibition, such as difficulty stopping a pre-planned motor action, and set shifting, which manifests itself in difficulty adjusting to new tasks or rules. One argument is that weaknesses in executive function contribute to the maintenance of stuttering and that the breakdown of language production could weaken executive function skills (Anderson & Ofoe, 2019; Ofoe et al., 2025). Nevertheless, these impairments are a challenge for current stuttering theories to explain, as some have argued that the executive function problems are lexical or phonological as opposed to general cognitive impairments (Bahrami et al., 2014; Lescht et al., 2024). This argument is supported by the findings from verbal fluency tasks in which individuals must verbally generate as many items conforming to specified rules within a set time limit. Bahrami et al. (2014) found that CWS and adolescents who stutter had lower scores in phonemic fluency (e.g., naming as many words as possible starting with F) than neurotypical adults but the authors observed that there were no group differences in semantic fluency (e.g., name as many words belonging to the category such as animals). Lescht et al. (2024) replicated these findings in both oral and typed production in AWS. Since both these types of verbal fluency tasks require articulation, this dichotomy suggests that speech-motor challenges cannot be the sole contributor to the stuttering traits. The lower scores on phonemic fluency tasks are thought to reflect difficulties in efficiently accessing and retrieving words from the mental lexicon, which is a core component of phonological and lexical processing. Ning et al. (2017) argued that AWS initiates speech motor preparation prematurely before the articulatory program is fully ready. This premature preparation could therefore cause a mismatch between motor commands and linguistic input, leading to delays in word access or stimulus processing. This is in line with EXPLAN theory, which proposes that stuttering results from a timing mismatch between independent linguistic planning and motor execution processes (Howell & Au-Yeung, 2002). This suggests that inhibitory control issues may be specific to the interference between the language and production systems. Given these findings, there is a clear need to disentangle core EF impairments from phonological processing impairments in PWS.

Since dyslexia and stuttering are two different neurodivergent conditions that influence different aspects of language processing, it might be expected that they should not co-occur in the same individuals. However, Elsherif et al. (2021) observed that dyslexia and stuttering co-occur at a rate much greater than expected by chance, with estimates of the degree of overlap being 30-40% (Elsherif et al., 2021). Despite the establishment of this empirical link between the two neurodivergent conditions, the nature of the co-occurrence remains unknown. In a subsequent experiment, Elsherif et al. (2021) used a suite of individual difference measures, including tests of orthography, phonology and semantics to disentangle the nature of this co-occurrence. Dyslexic adults and AWS were found to show similar levels of performance for measures of phonology only. The authors concluded that the two neurodivergent groups shared a core phonological processing impairment.

Stroop interference and lexical quality

As mentioned above, the Stroop task is a widely used measure of inhibitory control, competition and automaticity in reading (Stroop, 1935, see reviews by Parris et al., 2021, Scarpina & Tahini, 2017; Uttl & Graf, 1997). The task requires participants to ignore a certain dimension of a word (such as its form or meaning) and, instead, focus on responding to another particular dimension of a word (such as the colour of the font in which the word is presented), thereby requiring the participant to withhold habitual (or prepotent) responses and select appropriate conflicting, subdominant responses (McKenna & Sharma, 1995). This manipulation produces the Stroop interference effect, wherein incongruent trials (e.g., the word "RED" presented in a blue font) take longer for participants to name out loud than congruent trials where the two stimulus dimensions agree (e.g., the word "BLUE" displayed in a blue font; e.g., Dalrymple-Alford, 1972; Hasshim & Parris, 2021; see a review by MacLeod, 1991). In the incongruent condition, the participant must suppress the automatic word reading response to name the font colour. The magnitude of the resulting interference effect depends upon the degree of competition between the relevant (colour) and the irrelevant (word form or meaning) dimensions (Klein, 1964). Processing the colour dimension involves initial visual encoding, semantic activation, and phonological encoding for response preparation. In the naming Stroop (which requires vocal responses), this process is disrupted by competition at multiple levels: semantic interference from a conflicting word meaning, and phonological competition between word forms (e.g., Burca et al., 2021, 2022; Coltheart et al., 1999; Monsell et al., 2001; Parris et al., 2019, 2021). These complexities demonstrate the multifaceted nature of inhibitory control and the interference effects manifested in the Stroop interference task (see reviews by Parris et al., 2021; Scarpina & Tahini, 2017; Uttl & Graf, 1997). Performance on the Stroop task therefore has a strong verbal component. It involves the low-level processes involved in reading, which include paying attention to the presented information, recognizing visual patterns, connecting visual information with sounds, inhibiting the lexical-semantic representation of the word itself to name the font colour, and finally, being able to say the font colour (e.g., LaBerge & Samuels, 1974; Logan, 1997; Macleod, 1991; Mano et al., 2016). Stroop task performance strongly predicts text reading fluency and spelling (e.g., Booth et al., 2014; Mano et al., 2016; Megherbi et al., 2018). However, if the Stroop task strongly predicts text reading fluency and spelling, then the Stroop interference should be observed to be stronger in the dyslexic population, given their difficulties with literacy-related tasks.

Yet the findings concerning the Stroop interference effect in the dyslexic population are, in fact, equivocal. Some studies have reported that dyslexic people show larger Stroop interference effects than neurotypical readers (e.g., Stroop; Everatt et al., 1997; Faccioli et al., 2008; Helland & Asbjornsen, 2000; Protopapas et al., 2007, for younger teenagers; Kapoula et al., 2010, for older teenagers; Proulx & Emlary, 2015, for adults Go/No-Go task: McLean et al., 2011), others have failed to show similar differences in either children or adults (Reiter et al., 2005). One explanation for the mixed findings is that the dyslexia-related group differences on any given cognitive task may be the result of reduced or suboptimal reading experience (Huettig et al., 2018). In general, however, most studies on dyslexia have not used a measure of reading experience (e.g., the Author Recognition test; Huettig et al., 2018) to address this issue. It is important to note that, while we lack conclusive evidence concerning this bidirectional relationship between dyslexia and reading experience, findings have shown that the acquisition of reading and writing abilities demonstrably improves phonological awareness (e.g., Morais et al., 1986), verbal memory (e.g., Demoulin & Kolinsky, 2016), and spoken language (e.g., Huettig & Pickering, 2019). Therefore, including, and controlling for, reading experience can help to isolate its influence on cognitive tasks such as the Stroop. In addition, dyslexic adults and AWS perform similarly on phonological processing tasks (Elsherif et al., 2021) but AWS have no challenges in literacy. Therefore, by including this additional group as a means of comparison, we can assess whether the group differences between AWS and dyslexic adults shown on the naming Stroop task are due to varying levels of reading experience or a genuine inhibitory control impairment. However, in contrast to the relatively large corpus of research on Stroop performance in dyslexic individuals, there is no direct evidence relating to traditional naming Stroop performance in either CWS or AWS (but see, Hennessey et al., 2014; van Lieshout et al., 2014 for emotional stroop tasks).

Lexical and phonological processing impairments may therefore be fundamental to the increased Stroop interference effect observed in both dyslexic readers and PWS compared with that manifested by neurotypical readers. According to the lexical quality hypothesis (LQH; Perfetti, 1992, 2007, 2017; Perfetti & Hart, 2001, 2002), two key aspects of an individual's lexical representations are precision and redundancy. Lexical precision refers to the level of detail in orthographic (spelling) and phonological (sound) representations, as well as their connections to semantic representations, while redundancy describes the link between spelling, sound and semantics. Neurotypical readers typically have detailed orthographic and phonological representations that are strongly connected to semantic representations (see review by Perfetti, 2007, 2019). Consequently, such readers, on average, have faster access to semantic representations and exhibit more effective suppression of word reading than neurodivergent populations (Andrews, 2012, 2015; Andrews & Bond, 2009; Andrews & Hersch, 2010; Elsherif et al., 2022, 2023). This, in turn, allows their word recognition processes to proceed more smoothly, thereby facilitating the suppression of competing information and the monitoring of responses by the executive system. As a result, neurotypical adults are more likely to articulate the font colour more easily, due to more effective suppression mechanisms, leading to faster colour naming and a smaller Stroop interference effect. Executive control abilities enable the smooth parallel processing of word recognition, which facilitates the suppression of competing information, the monitoring of responses, and the articulation of the font colour, leading to faster colour naming and a smaller Stroop interference effect (Cohen et al., 1990). However, dyslexic individuals have less consistent links between a word's written and spoken form, resulting in lower quality lexical representations (Dobo et al., 2022; Perfetti, 2007). Because dyslexic readers take longer to access conceptual representations and are slower to suppress competing information, their speed and fluency in word identification are ultimately impaired. If these processes take longer to complete, less time remains for engaging executive processes to suppress the written word and name the font colour.

In skilled readers, the fast and fluent identification of words allows for more efficient inhibitory control. Efficient inhibitory control is critical for suppressing this automatic but irrelevant word-reading response (e.g., ignoring the printed word “red” when the font colour is blue). When skilled readers can quickly inhibit the word-reading task, cognitive resources can be re-allocated to enable faster and more accurate retrieval of the correct conceptual representation — in this case, the name of the font colour - from the mental lexicon. By quickly setting aside the automatic word-reading response, the individual can focus on the colour naming task, leading to more rapid and accurate performance. This, in turn, results in smaller Stroop interference. Indeed, more skilled readers have been found to perform better than less skilled readers on visual word recognition and visual word naming (Andrews & Hersch, 2010; Andrews & Lo, 2012; Elsherif et al., 2022, 2023), sentence probe memory tasks (Andrews & Bond, 2009), phonemic fluency tasks (e.g., Smith-Spark et al., 2017; Varvara et al., 2014) and semantic fluency tasks (e.g., Hall et al., 2017; Varvara et al., 2014). Essentially, the more frequently that an individual engages in reading- or speech-related activities, the richer their vocabulary becomes, along with an increasing richness in the details of sound (phonology) and meaning (semantics) associated with each word. These sharper lexical representations (Huettig & Pickering, 2019; Perfetti, 2007) would then help to suppress competing responses, such as the task-irrelevant word reading, which consequently facilitates the correct colour-naming responses and leads to a smaller Stroop interference effect.

The present Study

The aim of the current study is to investigate the contribution of lexical precision to Stroop performance in PWD and PWS. By comparing the performance of different participant groups (dyslexic adults, AWS, and neurotypical adults), we seek to explore how reading or speech difficulties interact with the suppression of competing responses in the Stroop task paradigm. We combine a group comparison with detailed individual difference measures of lexical and phonological processing and reading experience. In general, we predict that incongruent stimuli will take longer to name than congruent stimuli, thus yielding a Stroop interference effect consistent with prior research (e.g., Macleod, 1991).

If a significant proportion of Stroop interference were to be the result of poor phonological abilities then one would expect both dyslexic adults and AWS to show a similar degree of heightened Stroop interference relative to neurotypical readers. This is because a core feature of both dyslexia and stuttering is a phonological impairment (e.g., Anderson et al., 2006; Elsherif et al., 2021; Vellutino et al., 2004), which weakens the link between an abstract concept (like "redness") and its sound representation (the word "red"). In contrast, if the source of Stroop interference were to be the result of a lack of (or reduced) reading experience, then dyslexic adults should show stronger interference than neurotypical adults and AWS.

Moreover, individual differences in lexical precision should also modulate Stroop interference within these groups. Individual differences in language processing are multifaceted, encompassing phonology, orthography, and semantics. To better understand the relationship between the Stroop effect and these broader language processing abilities, we employ a comprehensive set of individual difference measures to assess specific components of phonology, orthography, and semantics that might modulate lexical retrieval. Although skilled readers generally outperform less skilled readers on various reading-related measures, isolating the specific component(s) underlying the Stroop effect is challenging due to the overlap in these skills (e.g., Acheson et al., 2008; Burt & Fury, 2000; Huettig et al., 2018; Martin-Chang & Gould, 2008). To disentangle these factors, we use Principal Component Analysis (PCA) to identify the underlying cognitive dimensions that explain the variance in our individual difference measures. In practice, we use a battery of individual differences tasks, including measures of orthography, phonology, semantics, and reading fluency. From this, PCA can empirically derive a smaller set of latent, uncorrelated components that best account for the overall variance in our dataset. For example, if orthography and phonology measures were highly correlated, they might load onto the same component. The resulting components, which represent the underlying cognitive dimensions, are then used to assess their unique contributions to Stroop interference in dyslexic readers and AWS.

Building on previous research (e.g., Everatt et al., 1997; Protopapas et al., 2007) that suggests executive control impairments in dyslexia, we predict a larger Stroop interference effect in dyslexic readers. We note, however, that the evidence for this finding is not entirely consistent across studies (Reiter et al., 2005). If this larger Stroop interference in dyslexic readers is due to reading-related or more general automaticity deficits (see Nicolson & Fawcett, 1990; Smith-Spark & Gordon, 2022), which implicate slower lexical access or difficulty suppressing competing responses, then it should not be observed in other populations who do not show reading difficulties, such as PWS. Conversely, if the Stroop interference effect stems from challenges with lexical precision, we should observe a similar interference effect in AWS, as they also experience difficulties with lexical precision (e.g., Elsherif et al., 2021), despite not having reading difficulties.

We also predict that the resulting construct of lexical precision will moderate the Stroop interference across groups (e.g., dyslexia: Proulx & Elmasry, 2015; stuttering: Maxfield, 2021). Specifically, we expect that the Stroop interference effect will be smaller for these individuals that score higher on the lexical precision component. Nevertheless, one key argument raised in the literature is that both neurodivergent groups share a phonological impairment (Elsherif et al., 2021) and, as a result, this leads to the prediction that a measure of phonological precision should moderate the Stroop interference manifested in both neurodivergent groups, while lexical and phonological precision will matter less for the neurotypical group because the connections between these constructs are more robust (Perfetti, 2007, 2019).

Method

Participants

Originally, 91 monolingual British undergraduate students without language impairment aged 18-22 (M=20.1±1.1 years) from the University of Birmingham participated in the study in return for course credits or monetary remuneration. Seven participants were excluded due to attrition and a further two as a result of the exclusion criteria used to define dyslexia being applied (see below). Further to this, 50 dyslexic monolingual adults aged 18-32 years (M=20.7±2.7 years; 27 males) were recruited from local universities, and 30 monolingual adults who stutter (AWS) aged 18-48 years (M=29.5±8.91 years; 22 males) were recruited from the general public for remuneration. The participant groups were matched as closely as possible in terms of age, educational level, bilingualism and handedness (see Table 1). Mirroring prevalence rates in the general population (e.g., Guitar, 2013; Hulme & Snowling, 2009), there were more male than female participants in both the dyslexia and stuttering groups. All dyslexic adults provided the researcher with a diagnostic assessment report documenting a childhood history of dyslexia and had no history of attention deficit hyperactivity disorder or a current diagnosis of persistent stuttering using the DSM-V criteria. In addition, none of the PWS had a history of attention deficit hyperactivity disorder or a current diagnosis of dyslexia using DSM-V criteria. The study was conducted in accordance with the British Psychological Society's ethical guidelines and approved by the first author's University's ethical committee. All participants had normal or corrected-to-normal vision and signed a consent form. The study was conducted in person.

Education level was recorded as the highest level of completed academic qualification declared by the participant and was classified as follows: UK A-levels (coded as 1), Bachelor's degree (coded as 2), Master's degree (coded as 3) and PhD (coded as 4) . The dyslexia and neurotypical groups did not differ significantly from each other in terms of average education level ($F(2, 161) = 2.09$, $p = .13$) or bilingualism ($F(2, 161) = 2.46$, $p = .08$) but there was a significant difference in age ($F(2, 161) = 57.87$, $p < .001$), gender ($F(2, 161) = 39.67$, $p < .001$) and handedness ($F(2, 161) = 6.8$, $p = .014$) such that there were more males and left-handed individuals in the PWD and PWS groups than in the neurotypical group, while the PWS were older than the PWD and neurotypical groups. PWD had more left-handers than PWS.

Table 1.

Mean and standard deviation for demographic variables for all diagnostic groups.

	Neurotypical (N = 84)	**Dyslexia (N = 50)**	**Stuttering (N = 30)**
	Mean (*SD*)	**Mean (*SD*)**	**Mean (*SD*)**
Age (years)	20.18 (1.03)	20.72 (2.7)	29.47 (8.91)
Education	1.04 (0.23)	1.14 (0.35)	1.13 (0.43)
Gender (%female)	91.7 (0.28)	46 (0.50)	27 (0.42)
Handedness (%Right)	89 (0.41)	62 (0.49)	87 (0.33)
Bilingualism (%Bilingual)	0 (0)	4 (0.20)	7 (0.25)

Apparatus

Stimuli were presented on a standard computer monitor and responses were made on a QWERTY keyboard. Stimulus presentation and response timing were controlled by E-Prime (Psychology Software Tools, 2002).

Design

The experiment had a 2 (congruent: congruent vs. incongruent) x 3 (Group: neurotypical vs. dyslexia vs. stuttering) mixed factor design. Reaction times (RT) for correct responses only served as the dependent variable of interest. The Stroop interference effect is calculated as the reaction times in the congruent condition subtracted from the reaction times in the incongruent condition.

Procedure

At the start of each experiment, demographic information (gender, age, handedness) and informed consent was collected. Participants were informed that they would perform a naming Stroop and a suite of individual difference tasks and two other tasks (unrelated to the current experiment). Participants sat approximately 60 cm from the screen.

General Procedures for the tests. Each participant completed all the components over three sessions. Each session lasted approximately an hour. All participants completed the tests in the same order (see Figure 1; Elsherif et al., 2021, 2022, 2023). Participants were assessed on several measures of orthography, phonology, reading fluency, semantics, non-verbal intelligence and inhibitory control (see the Stroop section below), which are described in detail in Elsherif et al. (2021, 2022, 2023; see Table 1).

Table 1.

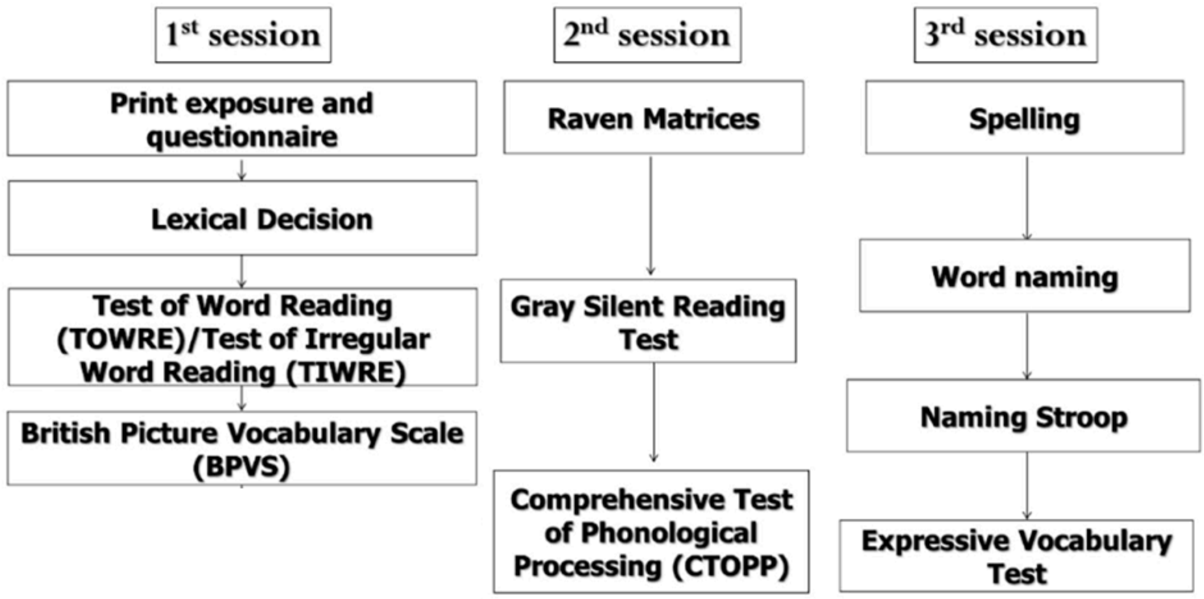
The individual difference measures used in the current experiment and their groupings. Reproduced from Elsherif, M. M., Wheeldon, L. R., & Frisson, S. (2021). Phonological precision for word recognition in skilled readers. *Quarterly Journal of Experimental Psychology, 75*(6), 1021-1040. https://doi.org/10.1177/17470218211046350 under CC BY-NC 4.0 license.

Tests	**Administration**	**Measures**
	Orthography	
Author, title and magazine recognition test[a]	Mark known authors, book titles and magazine, respectively	Print exposure
Spelling[b]	Spell the word dictated	Spelling
	Phonology	
Phoneme elision[c]	Remove a phoneme from a real word to form a new word	Phonological awareness
Memory for digits[c]	Recall numbers in the same order	Phonological working memory
Nonword repetition[c]	Repeat nonwords	Phonological working memory
Phoneme reversal[c]	Reversal of pseudowords to form an existing word	Phonological processing
Rapid letter naming[c]	Read letters as fast as you can	Grapheme-phoneme conversion
	Reading Fluency	
TOWRE: Sight word efficiency[d]	Read words for 45s	Word decoding
TOWRE: Phoneme decoding[d]	Read pseudowords for 45s	Phonological decoding
TIWRE: Irregular word reading efficiency[e]	Read irregular words	Lexical reading
	Semantics	
Expressive vocabulary test[f]	Answer the question in relation to the picture	Expressive vocabulary
British Picture Vocabulary Scale[g]	Choose out of 4 pictures that reflect the word said	Receptive vocabulary
Gray silent reading test[h]	Read stories and answer questions	Comprehension
Raven's standard progressive matrices[i]	Fit the overall patterns with missing panels	Non-verbal intelligence

Note: Test of Word Reading (TOWRE) and Test of Irregular Word Reading Efficiency (TIWRE). [a]Cunningham and Stanovich (1990) and Stanovich and West (1989); [b]Elliott et al. (1996), [c]Wagner et al. (1999), [d]Torgesen et al. (1999), [e]Reynolds and Kamphaus (2007), [f]Williams (2007), [g]Dunn et al. (1997), [h]Wiederholt & Blalock (2000), [i]Raven (1960) and [j]Stroop (1935).

Figure 1.

An overview of the three experimental sessions. Reproduced from Elsherif, M. M., Wheeldon, L. R., & Frisson, S. (2021). Phonological precision for word recognition in skilled readers. *Quarterly Journal of Experimental Psychology, 75*(6), 1021-1040. https://doi.org/10.1177/17470218211046350 under CC BY-NC 4.0 license.



Standardized tests

The same standardized tests were used as Elsherif et al. (2021, 2022, 2023). The experiment was conducted in-person.

Demographic questionnaire

Participant demographic data, including age, gender, and handedness, was gathered via a questionnaire.

Language measures

Phonological processing

Phonological processing abilities were assessed using five subtests from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999), namely Phoneme Elision, Rapid Letter Naming (RLN), Memory for Digits, Non-word Repetition, and Phoneme Reversal. The Phoneme Elision task, consisting of 18 items, required participants to orally remove a specified phoneme from a spoken word and then articulate the remaining word (e.g., producing "at" after deleting /m/ from "mat"). Non-word Repetition similarly comprised 18 items, during which participants repeated auditorily presented non-words. For the 18-item Phoneme Reversal task, participants rearranged the phonemes of a given word or non-word to form a real word (e.g., "na" becoming "an"). For these three subtests, the total number of correct responses was recorded. The Memory for Digits subtest, consisting of 21 items, evaluated the ability to recall and repeat spoken digit sequences, with the score based on the number of correctly recalled sequences. Lastly, Rapid Letter Naming (RLN) measured the time taken to name 36 printed letters and was the only one of the phonological processing tasks not to have a time limit.

Reading fluency

Reading fluency was assessed using subtests from the Test of Word Reading Efficiency (TOWRE; Torgesen, Rashotte, & Wagner, 1999): the Sight Word Efficiency subtest (TOWRE-W), which measures sight word reading, and the Phonemic Decoding Efficiency subtest (TOWRE-P), which assesses phonemic decoding through nonword reading. Additionally, irregular word reading was evaluated using the Test of Irregular Word Reading Efficiency (TIWRE; Reynolds & Kamphaus, 2007). For the TOWRE-W, participants read aloud a list of 108 regular words (e.g., "strange") within a 45-second time limit. The TOWRE-P followed the same format but presented 66 pronounceable nonwords (e.g., "tindor"). While standard administration for both neurotypical and dyslexic adult groups involved a 45-second time limit, AWS were given an extended time limit of three minutes to account for their known difficulties with time pressure (Perkins et al., 1991). However, to ensure consistent scoring across all participants, only responses completed within the initial

45 seconds were included in the analysis. One AWS participant's disfluent responses were excluded, but this exclusion did not affect the overall findings. The TIWRE required participants to read 25 irregularly spelled words (e.g., "great") without a time constraint. Nonword reading primarily assesses phonological decoding (grapheme-phoneme conversion), whereas irregular word reading measures access to orthographic representations (Bowey & Rutherford, 2007; Cortese & Simpson, 2000).

Print exposure

Given that print exposure has been linked to orthographic and phonological processing efficiency in university students (Chateau & Jared, 2000), and its documented influence in dyslexia research (Huettig et al., 2018), we included a measure of print exposure to control for its potential confounding effects on performance. Our measure of print exposure was based on adapted versions of two well-established recognition tests: the Author Recognition Test (ART; Stanovich & West, 1989) and the Title Recognition Test (TRT; Cunningham & Stanovich, 1990). Each test presented participants with a checklist of 100 items, equally divided between real items and foils (50 real, 50 foils). Participants indicated their familiarity with each item without a time limit. The ART focused on author names, while the TRT assessed recognition of book titles, including novels, plays, and poetry. For our British sample, we updated the original item lists to include a mix of classic and contemporary authors and titles, partly by consulting Amazon's top 100 lists. Pre-testing with 100 participants from the same population confirmed that responses were normally distributed. To calculate print exposure scores, we subtracted false alarms (incorrectly identified foils) from hits (correctly identified real items) for both the ART and TRT. A composite print exposure score was then derived by averaging the ART and TRT scores.

Receptive vocabulary

Receptive vocabulary was evaluated using the British Picture Vocabulary Scale, Second Edition (BPVS-II; Dunn, Dunn, Whetton, & Burley, 1997). During this assessment, participants heard a recorded word and were required to select the corresponding image from a set of four options. Six specific sets of vocabulary items (Sets 9–14) were administered, with no time constraints. Each participant's total number of correct responses was recorded.

Reading comprehension

Reading comprehension was evaluated using stories 4 through 9 of the Gray Silent Reading Test (GSRT; Widerholt & Blalock, 2000). Participants read six stories, which increased in complexity, at their own pace. Following each story, they answered five multiple-choice questions. The total number of correct answers summed across stories was recorded as their score.

Spelling

Spelling ability was evaluated using a 20-word dictation subtest, an adaptation from the British Ability Scale (Elliott, Smith, & McCulloch, 1996). This measure was included as spelling is considered an indicator of lexical precision (Andrews & Bond, 2009; Andrews & Hersch, 2010). Participants had unlimited time to spell each word, and their performance was scored based on the number of correct spellings.

Non-verbal intelligence measure

Raven

Non-verbal intelligence was assessed with the Raven’s Standard Matrices (Raven, 1960). Participants completed 60 pattern-completion problems of increasing difficulty at their own pace. The total number of correct responses was recorded. Unlike prior research with aging populations that imposed time limits, the current study removed this constraint to isolate non-verbal intelligence from processing speed effects.

Stroop task

The experimental stimuli for the vocal Stroop task were five colour words ('red', 'blue', 'green', 'yellow' and 'purple') presented in either red, blue, green, yellow or purple coloured font. Stimuli were presented against a grey background in white font (Arial, size: 34). Each participant received five blocks of twenty trials each. Participants were presented with 75 congruent trials and 25 incongruent trials. By presenting a higher proportion of congruent trials (e.g., the word "RED" printed in red ink), participants develop a stronger expectation for congruence, thus increasing the strength of the Stroop interference effect (see reviews by Algom & Chajut, 2019; Macleod, 1991). Within the experiment, the trials were presented in an individually randomised sequence. Short breaks were inserted after 25 trials of the task; at this point, participants could continue the experiment by pressing the space key.

Before the experimental trials began, participants familiarised themselves with the task in a short practice block comprising four trials. The stimuli for the practice block were randomly chosen from the set of word-colour combinations. Trials in both the practice and test block were randomly ordered for each participant. Each trial started with a white fixation cross that was presented for 500 ms in the centre of the screen. The fixation cross was replaced by one of the coloured word stimuli, with participants being instructed to name the font colour out loud and to ignore the naming word (Stroop, 1935). They did so by speaking into a microphone connected to a Sony DAT recorder (PCM-M1), enabling future offline analysis of the naming data. This presentation was followed by an inter-trial interval of 1000ms. This task lasted for approximately five minutes.

Analysis

Calculation of composite scores and correlations

For a Principal Component Analysis (PCA) to be reliable, five to 10 participants are typically required for each variable entered (Kass & Tinsley, 1979). We reduced the overall number of variables (i.e. 15) to eight variables due to limitations in the size of our participant pool (i.e. N = 164) and to avoid redundancy between measures due to high correlations (see examples by Andrews & Lo, 2012; Elsherif et al., 2021, 2022, 2023; Holmes et al., 2014). The procedure was as follows: We first z-transformed the scores for the individual difference measures to ensure different scores could be directly compared. Following this transformation, we conducted Pearson's product-moment correlations (*r*) to identify related measures. Specifically, this step was undertaken to confirm that the individual measures intended to be grouped together for subsequent averaging were indeed sufficiently intercorrelated, thereby supporting the assumption that they reflected the same underlying construct. Strong correlations between these measures would indicate that they were assessing the same underlying construct (e.g., nonword repetition and memory for digits are reflective of the phonological working memory construct; Wagner et al., 1999). Establishing these strong intercorrelations *before* averaging their respective z-scores was crucial to ensure that the combined composite scores genuinely represented a unified construct rather than an arbitrary amalgamation of disparate measures. We therefore averaged the z-scores of these related measures to create a composite score. Afterwards, we subjected the remaining individual difference measures and composite scores to identify underlying components. This approach reduced the number of variables entered into the PCA, while preserving the underlying information for further analysis.

Results

Demographic variables, attrition, and cognitive and language tests

Our participants were homogeneous in their demographics. However, there was an appropriate level of variability in scores on all of the tests across groups (see Table 2).

Table 2.
Means and standard deviation of all measures[4].

Measure	Neurotypical (*n* = 84)	Dyslexia (n = 50)	Stuttering (n = 30)
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Author Recognition Test (out of 50)	15.07 (7.69)	16.54 (9.61)	26.23 (7.69)
Title Recognition Test (out of 50)	18.49 (6.13)	19.76 (8.82)	23.8 (11.63)
Magazine Recognition Test (out of 50)	11.26 (4.60)	11.26 (4.60)	11.26 (4.60)
British Picture Vocabulary Scale (out of 60)	42.74 (7.27)	42.40 (7.78)	39.60 (9.93)
Expressive Vocabulary Test (out of 118)	71.86 (8.01)	72.95 (10.36)	77.80 (13.93)
TOWRE Sight Word Efficiency (out of 108)	87.34 (11.24)	78.82 (12.10)	73.73 (20.10)
TOWRE Phoneme Decoding (out of 65)	57.88 (5.56)	49.3 (7.15)	51.87 (10.94)
TIWRE (out of 25)	21.19 (1.86)	20.26 (2.25)	21.57 (2.88)
CTOPP Phoneme Elision (out of 20)	16.67 (2.38)	15.32 (3.04)	15.07 (3.60)
CTOPP Memory for Digits (out of 21)	16.65 (2.14)	14.34 (2.04)	15.33 (3.08)
CTOPP Non-Word Repetition (out of 18)	13.68 (1.73)	10.68 (2.23)	9.9 (2.90)
CTOPP Rapid Letter Naming (ms)	26.27 (4.79)	32.59 (9.42)	42.13 (52.94)
CTOPP Phoneme Reversal (out of 18)	11.43 (2.62)	8.96 (2.49)	9.93 (3.83)
Gray Silent Reading (out of 30)	22.30 (3.30)	21.18 (4.05)	23.27 (4.53)
Spelling (out of 20)	16.49 (2.37)	13.24 (2.95)	14.87 (4.19)

Note. CTOPP = Comprehensive Test of Phonological Processing, TOWRE = Test of Word Reading Efficiency, TIWRE = Test of Irregular Word Reading Efficiency.

Correlation

To conduct the PCA, we began with Pearson r correlation, as we combined two highly correlated vocabulary measures (r = .61) into a single "CVocab" score by averaging their standardized values. This provides a more comprehensive picture of vocabulary ability.

Similarly, we created a "CMemory" score by averaging the standardized scores from two highly correlated measures of phonological working memory (nonword repetition and digit memory; $r = .53$). Three reading fluency measures were combined into a "CReadingFluency" score. These measures correlated (TOWRE word reading & Rapid Letter Naming: $r = .48$; TOWRE phonemic decoding & Rapid Letter Naming: $r = -.54$; TOWRE word reading & phonemic decoding: $r = .67$) and averaging their standardized scores provided a more well-rounded assessment of reading fluency. Finally, a "ZPrintexposure" score was created by averaging standardized scores from two strongly related measures of print exposure (ART and TRT; $r = .84$). See supplementary material 1 for details on the correlation and the original PCA without the variables being combined.

Table 3 shows the correlations between these composite scores and other individual differences measures. These correlations align with existing research (e.g., Martin-Chang & Gould, 2008 for print exposure and vocabulary; Burt & Fury, 2000 for print exposure and spelling). Importantly, many individual difference measures exhibited high collinearity (correlations ≥ .3). This redundancy justifies using a multivariate approach like PCA for further analysis.

Table 3.

Correlations between tasks.

	ZVocab	PE	ZMemory	ZRF	PR	TIWRE	Spell	GSRT
PE	0.22**							
ZMemory	0.33**	0.23**						
ZRF	0.12	0.16	0.34**					
PR	0.32**	0.41***	0.48***	0.27**				
TIWRE	0.48***	0.31**	0.35***	0.06	0.37***			
Spell	0.44***	0.34**	0.46***	0.28**	0.47***	0.50***		
GSRT	0.53***	0.23**	0.34**	0.05	0.24**	0.40***	0.40***	
ZPE	0.59***	0.05	0.11	-0.11	0.08	0.47***	0.30**	0.44***

Note. Standard vocabulary composite measure (ZVocab), Phoneme Elision (PE), Standard phonological working memory composite measure (ZMemory), Standard reading fluency composite measure (ZRF), Phoneme reversal (PR), Gray Silent Reading Test (GSRT), Test of Irregular Word Reading Efficiency (TIWRE), Standard print exposure composite measure (ZPE). **, p < .01 and *** p < .001.

Principal Component Analysis

To understand the relationships between the different reading skills that we measured (Table 3), we performed a Principal Component Analysis (PCA). This statistical technique helps us identify underlying factors (components) that explain most of the variability in the data. Before running the PCA, we checked if our data was appropriate for this analysis, using the software package GPA rotation (Bernaards & Jennrich, 2005), within the R statistical programming open code software (R Development Core Team, 2017). The Kaiser-Meyer-Olkin (KMO) measure was .82, which is well above the recommended value of .5 (Field, 2009), indicating good sampling adequacy. Bartlett's test of sphericity was significant ($p < .001$), showing that the correlations between the variables were appropriate for PCA.

We conducted the PCA using both varimax (orthogonal) and oblique rotations. Varimax rotation assumes the components are independent, while oblique rotation allows for some correlation between them. We focused on the varimax results (Table 4) because the highest correlation between components from both rotations was only .26, suggesting minimal overlap (Tabachnick et al., 2007). Based on the loadings, these two components were assigned construct names indicative of their component variables and are listed in the order of variance explained in Table 4. The initial analysis revealed two main components that explained 52% of the variance in reading behaviors. Each component groups variables that demonstrate high absolute loadings (greater than 0.45) on that component, thereby indicating that these variables share a common underlying factor. The loadings themselves can be either positive or negative, reflecting the direction of their relationship with the underlying factor. Components show positive or negative loadings. Positive loadings give inclusionary criteria and describe the underlying construct of the component. Negative loadings provide exclusionary criteria and show an inverse relationship to the construct of the component.

The first component, accounting for the most variance, includes measures of vocabulary, print exposure, spelling, and reading comprehension (all positive loadings). The larger scores might benefit from stronger and richer resonance between the sublexical representations, allowing more efficient bottom-up processing, consistent with the term of lexical precision (Perfetti, 2007). Put simply, lexical precision is the connection between lexical orthographic representation to higher-level semantic knowledge, in spite of vocabulary measures not being orthographic in nature.

To test for group differences between the lexical precision component, we computed a score on this component for each participant and conducted independent samples *t*-tests (in line with the approach taken by Henderson et al., 2023). The AWS scored higher on the lexical precision component than neurotypical adults ($t(112) = 2.93$, $p = .004$) and dyslexic adults ($t(78) = 2.34$, $p = .02$). Dyslexic adults did not significantly differ from neurotypical adults in this measure ($t(133) = 0.38$, $p = .71$).

The second component groups measures of reading fluency, phonological awareness, working memory, and spelling. Positive loadings here suggest that strong phonological skills (working memory, awareness) contribute to efficient processing of sound-to-letter relationships during reading (Perfetti, 2007). This component emphasizes the role of phonological skills in accurate reading, thus reflecting a component of *phonological precision*.

To test for group differences between the phonological precision component, we computed a score on this component for each participant and conducted independent samples *t*-tests (in line with the approach taken by Henderson et al., 2023). The neurotypical adults scored higher on the phonological precision measure than dyslexic adults ($t(133) = 10.68$, $p < .001$) and people who stutter ($t(112) = 5.88$, $p < .001$). However, adults who stutter did not significantly differ from dyslexic adults ($t(78) = 0.53$, $p = .60$).

Table 4.

Components produced by the PCA.

Component 1 Lexical precision	Loading value	**Component 2** Phonological precision	Loading value
CPrint exposure	0.85	Phoneme reversal	0.75
CVocab	0.79	CMemory	0.74
GSRT	0.71	CRF	0.67
TIWRE	0.68	Spelling	0.62
Spelling	0.46	Phoneme Elision	0.56
% variance	*0.30*		*0.27*
Cumulative variance	*0.30*		*0.57*

Note: Standard vocabulary composite measure (ZVocab), Standard phonological working memory composite measure (ZMemory), Standard reading fluency composite measure (ZRF), Gray Silent Reading Test (GSRT), Test of Irregular Word Reading Efficiency (TIWRE), Standard print exposure composite measure (ZPE).

General linear mixed effect model (GLMM)

Following the PCA, we placed the resulting components into the General linear mixed model. We applied GLMM to the naming Stroop latencies, using the Ime4 package (Bates, Maechler & Dai, 2010) in R statistical programming open code software (R Development Core Team version 3.6.1., 2017). This approach accounts for both individual participant differences and variations across different stimuli. The model included congruency (sum coded with congruent as the intercept) as a fixed effect with all slopes and intercepts allowed to vary at random by subjects and items. In addition, the fixed factor of group was included as a between fixed effect (sum coded with neurotypical as the intercept).

To avoid overfitting the model and ensure reliable results, we initially used a maximal random structure, with a focus on reducing Type I and II errors. We included the fixed factor of congruency as part of the by-subject and by-trial random slopes (Barr et al., 2013). However, in our analysis, the maximal model failed to converge. This is a known issue (Winter, 2019). As a result, we simplified the random structure by removing random slopes until the model converged. For models, where even a simplified random structure did not converge, a non-random model was used (Veldre & Andrews, 2014). The two components identified from the PCA analysis (lexical and phonological precision) were included as additional fixed effects in the model and analyzed as continuous variables. If interactions arose between the PCA components and other factors, we transformed the continuous PCA data into binary categories ("high" vs. "low") for further analysis. The recoding was done by splitting the data from a variable into two sets so that the number of data points per set was as closely matched as possible.

We used "drop1" to identify the most parsimonious model (i.e., the least complex) by removing non-significant fixed effects until the Bayesian Information Criterion (BIC) reached its minimum value. Significance was determined by absolute *t*-values greater than 2 (α = .05). We additionally calculated Cohen's *d* effect sizes to quantify the magnitude of differences within groups. The calculation of ΔM was conducted with estimated marginal means and the standardisation of σ was computed with the total variance from covariance model estimates (Cohen, 1988; Westfall et al., 2014).

Outlier rejection

Following the methods of Loenneker et al. (2024), we pre-processed our data to ensure reliability and accuracy. First, we conducted a quality check on all participant data. We excluded trials with reaction times (RTs) under 250ms, as these typically indicate impulsive or anticipatory responses or responses being reflective of fillers (e.g., umms, hmms). A high number of such trials could suggest a lack of engagement with the task. After this step, we calculated the proportion of correct responses for each participant. Participants who performed at or below chance level (accuracy < 60%) were removed from the dataset; however, no participants met this exclusion criterion. Accuracy was high for all conditions and since none of the models using accuracy data reached convergence, we will not discuss accuracy further. Next, only correct trials were included in the RT analyses. Following the inclusion of only correct trials, we removed outlier RTs. Outliers were defined as any RTs that fell more than ±2.5 standard deviations from an individual participant's median, instead of the mean, RT, consistent with the approach of Berger and Kiefer (2021). We removed 7.3% of the data prior to the analyses due to the errors made and the outlier rejection used for the naming Stroop task. Average RTs, SDs and proportion of correct responses for each condition are shown in Table 5. Adults who stutter were slower than neurotypical individuals and dyslexic adults (see Table 6 for the output of this model). Dyslexic adults were slower than neurotypical adults, albeit the effect was not significant. Participants took less time to respond to congruent stimuli than to incongruent stimuli. The random model for the Stroop task converged (see Appendix C for the final model).

Table 5.

Mean response times and proportion correct for congruency conditions for neurotypical adults, dyslexic adults and AWS groups.

Congruency	Neurotypical	Dyslexia	Stuttering
Congruent			
RT	581 (127)	612 (145)	717 (254)
P correct	.99 (0.11)	.99 (0.09)	.99 (0.10)
Incongruent			
RT	699 (160)	764 (175)	890 (323)
P correct	.93 (0.26)	.92 (0.27)	.97 (0.17)
SI			
RT	-118	-152	-173
P correct	.06	.07	.02

Note. SI (Stroop Interference), response times (RT); Proportion (P); Response times are measured in milliseconds and standard deviations are in parentheses. Stroop interference was calculated as the difference score between Congruent and Incongruent conditions.

Table 6.

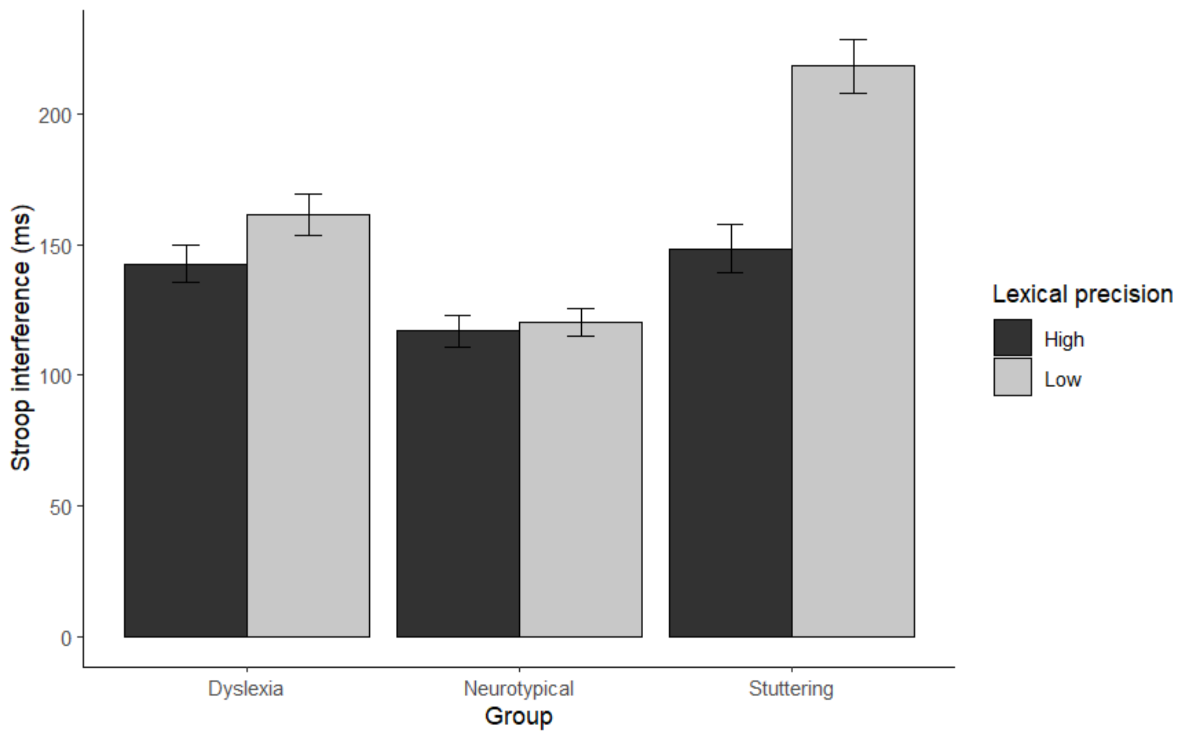
The minimal model output for RTs for Stroop.

Fixed Effects	Estimate	Std. Error	95% CI	95% UCI	*t* values
A: NT as intercept, dyslexia vs. stuttering					
(Intercept)	6.3371	0.0174	6.3030	6.3712	364.14*
Manipulation					
Congruent	0.1834	0.0009	0.1816	0.1852	21.36*
Group-dyslexia	0.0542	0.0280	-0.0007	0.1091	1.94 .
Group- stuttering	0.1534	0.0348	0.0852	0.2216	4.40*
Individual differences					
Lexical Precision	-0.0247	0.0254	-0.0745	0.0251	-0.98
Interactions					
group-dyslexia * congruent	0.0296	0.0139	0.0024	0.0093	2.12*
Group-stuttering * congruent	0.0531	0.0173	0.0192	-0.0203	3.06*
group-dyslexia * Lexical Precision	0.0448	0.0336	-0.0211	0.0749	1.33
group-stuttering * Lexical Precision	0.0842	0.0318	0.0219	-0.0111	2.65*
Congruent * lexical precision	0.0125	0.0126	-0.0122	0.0365	0.99
Group – dyslexia * congruent * lexical precision	-0.0371	0.0168	-0.07	0.1541	-2.21*
Group – stuttering * congruent * lexical precision	-0.0531	0.0158	-0.0841	0.1806	-3.35*
B: Dyslexia (intercept) vs. stuttering					
(Intercept)	6.3934	0.0249	6.3446	6.4422	267.05*
Manipulation					
Group	0.1072	0.0402	0.0284	0.1860	2.67*
Congruent	0.2194	0.0095	0.2008	0.2380	23.17*
Individual differences					
Lexical precision	0.0423	0.0157	0.0115	0.0731	2.69*
Interaction					
Congruent * Lexical precision	-0.0314	0.008	-0.0471	-0.0157	-4.14*

Note. NT = neurotypical; * $p < .05$.

Figure 1.

Reaction time (RT) Stroop interference effects (in ms) for dyslexic adults, AWS and neurotypical individuals and separated by the lexical precision composite. Error bars represent the 95% confidence interval for each condition. The Stroop interference effect is calculated as the mean reaction times in the congruent condition subtracted from the mean reaction times for the incongruent condition.



In the reduced model with dyslexic adults, AWS and neurotypical adults, there was a significant three-way interaction between group, congruency and lexical precision for neurotypical and dyslexic adults (Figure 1). Dyslexic adults with low lexical precision were more likely to show larger Stroop interference than those with high lexical precision, whereas there was little difference between neurotypical individuals in terms of lexical precision for Stroop interference. The full model was split into two sub-models: neurotypical and dyslexia. Group was removed from the equation and the same procedures from the full model for the analysis and random structure were applied to the sub-models. In the neurotypical submodel, there was no significant interaction between congruence and lexical precision (*b*=0.012, *t*=1.05, *p*=.30), nor an effect of lexical precision (*b* = -0.025, *t*=-1.06, *p*=.29). However, there was a significant effect of congruency (*b* = 0.18, *t*=23.31, *p*<.001), with congruent stimuli taking less time to name than incongruent stimuli. In the dyslexia submodel, the interaction between congruence and lexical precision was found to be significant (*b*=-0.02, *t*=-2.10, *p*=.04), such that dyslexic adults with low lexical precision were more likely to show larger Stroop interference than those with high lexical precision.

In addition, we observed a significant three-way interaction between group, congruence and lexical precision for neurotypical adults and AWS (Figure 1). People who stutter with low lexical precision were more likely to show larger Stroop interference than those with high lexical precision, whereas there was little difference between neurotypical individuals in terms of lexical precision for Stroop interference. The full model was split into two sub-models: neurotypical and stuttering. Group was removed from the equation and the same procedures from the full model for the analysis and random structure were applied to the sub-models. In the stuttering submodel, an interaction between congruence and lexical precision was found to be significant (*b*=-0.04, *t*=-3.73, *p*<.001) such that people who stutter and have low lexical precision were more likely to show larger Stroop interference than those with high lexical precision.

Split by lexical precision

The full model was also split into two sub-models: high lexical precision and low lexical precision. Lexical precision was removed from the equation and the same procedures from the full model for the analysis and random structure were applied to the sub-models. In the high lexical precision submodel, there was no significant interaction between congruence and group (*b* = 0.005, *t* = 0.21, *p* = .84). However, there was a significant effect of congruency (*b* = 0.18, *t* = 13.51, *p* < .001), with congruent stimuli taking less time to name than incongruent stimuli. In addition, there was a main effect of group (*b* = 0.10, *t* = 2.39, *p* =

.02) such that dyslexic adults took longer to name items than neurotypical adults. In the low lexical precision submodel, the interaction between congruence and group was found to be significant (b = 0.04, t = 2.19, p = .02), such that dyslexic adults with low lexical precision were more likely to show larger Stroop interference than neurotypical adults with low lexical precision.

The full model that had neurotypical adults, AWS and dyslexic adults was also split into two sub-models: high lexical precision and low lexical precision. Lexical precision was removed from the equation and the same procedures from the full model for the analysis and random structure were applied to the sub-models. In the high lexical precision submodel, there was no significant interaction between congruence and group (b = 0.005, t = 0.21, p = .84). However, there was a significant effect of congruency (b = 0.18, t = 13.51, p < .001), with congruent stimuli taking less time to name than incongruent stimuli. In addition, there was a main effect of group (b = 0.24, t = 5.10, p < .001) such that AWS took longer to name items than neurotypical adults. In the low lexical precision submodel, the interaction between congruence and group was found to be significant (b = 0.081, t= 3.09, p < .001), such that AWS with low lexical precision were more likely to show larger Stroop interference than neurotypical adults with low lexical precision.

In the reduced model between dyslexic adults and AWS, we observed that there was a significant interaction between congruence and lexical precision (Table 6). People who have high lexical precision showed smaller Stroop interference than those with low lexical precision. There was a significant effect of congruence in neurodivergent individuals with high lexical precision (b=0.196, t=14.88, p<.001) and neurodivergent individuals with low lexical precision (b=0.237, t =15.23, p< .001).

Discussion

Previous research has demonstrated phonological impairments in dyslexic children and adults (e.g., see review by Vellutino et al., 2004) and in CWS and AWS (e.g., Anderson & Wagovich, 2010; Elsherif et al., 2021). Building upon this work, the current study investigated a separate but related question: whether the suppression of competing responses is also moderated by lexical-semantic processes. This finding replicated previous research in finding elevated Stroop interference in dyslexia (Kapoula et al., 2010; Stubenrauch et al., 2013) and supported the notion of a lexical competition impairment in dyslexia (Goranova et al., 2021; Jones et al., 2016). In the case of AWS, this is the first study to employ the standard Stroop task. For this group we observed an elevated Stroop interference compared to neurotypical adults. Our findings align with more general arguments for a lexical selection impairment in AWS (Maxfield, 2021).

Further to corroborating previous findings, the current study extended previous research by providing a finer grained analysis of performance. We found that for both dyslexic adults and adults who stutter, Stroop interference was related to varying levels of lexical precision. Specifically, individuals in these groups who demonstrated higher levels of lexical precision were more likely to show a comparable level of Stroop interference to that shown by neurotypical adults. Conversely, those with lower levels of lexical precision tended to exhibit greater Stroop interference than neurotypical adults. We found no group differences between dyslexic adults and AWS concerning the magnitude of Stroop interference that they manifested. Although we must interpret this finding cautiously due to potential limitations in statistical power and more general caution around interpreting null results, these results align with theories (Andrews & Bond, 2009; Andrews & Hersch, 2010; Andrews & Lo, 2013; Perfetti, 2007, 2019) that emphasise the crucial role of lexical precision in word recognition. Proficient word processing, including tasks like the Stroop, relies on the rapid and accurate online integration of multiple lexical components—such as orthography, phonology, and semantics—to form high-quality lexical representations (Perfetti, 2007, 2019). If the competitors of the word are easily suppressed, then the word itself is easily suppressed, and the colour can thus be named more efficiently.

In addition, when neurotypical adults are removed from the model to directly compare the performance of AWS and dyslexic adults, we observed a pattern of lexical precision and congruence such that people with high lexical precision showed smaller Stroop interference than those with low lexical precision. Here, Stroop task performance, particularly the ability to efficiently resolve conflict, is not merely a measure of domain-general inhibitory control. Instead, it is deeply intertwined with the underlying quality and accessibility of these lexical representations (see Jones et al., 2010; Maxfield, 2021, for a concept access account of naming delays in dyslexia and stuttering; see also Smith-Spark & Moore, 2009, who observed a lack of age-of-acquisition effect in famous face naming, indicative of poor long-term semantic representation). While both groups experience challenges in aspects of language processing (e.g., phonology + fluency in both dyslexia and stuttering, plus spelling in dyslexia; Elsherif et al., 2021), our findings, *if taken in conjunction with the observed role of lexical precision in Stroop performance across groups*, suggest that their core impairment might *not* primarily manifest itself as a generalized *impairment in conflict resolution itself*. Instead, any observed differences in Stroop performance, or lack thereof, would seem to be more nuanced.

We observed a three-way interaction between lexical precision, group and congruency. Dyslexic adults who have lower levels of the lexical precision component showed larger Stroop interference than their neurotypical peers, while we observed those who have higher levels of the lexical precision component had comparable levels of Stroop interference to their neurotypical peers. This indicates that these groups, despite their distinct challenges, achieve a level of online integration of the Stroop's verbal components that may not result in a globally elevated interference effect, especially when considering the compensatory strategies often employed by these groups who may have high levels of print exposure to face alternative strategies to aid in the resolution of competition. Therefore, while their underlying processing mechanisms may differ, increased print exposure may help dyslexic adults and AWS to integrate the visual word and colour information, resulting in similar levels of overall Stroop interference to neurotypical controls. This suggests that while robust bottom-up lexical activation is a pre-requisite for both groups, those with higher lexical precision— regardless of neurodivergent status — are better equipped to resolve the conflict. This could be due to more efficient top-down control or the use of compensatory strategies that allow for rapid suppression of competing lexical information, ultimately reducing the global interference effect to name the colour (Andrews & Hersch, 2010; Elsherif et al., 2022, 2023; Perfetti, 2007).

The current study provides novel insights into the interplay between inhibitory control, as measured by the Stroop task, and lexical-semantic processing in dyslexic adults and AWS. Our findings extend existing theories of reading and speech production such as the LQH by demonstrating that lexical precision moderates Stroop interference in both groups. The ability to resolve conflict in language-related tasks, such as suppressing the printed word name in the Stroop task to name the font colour, is not solely a measure of generalised inhibitory control, but is deeply intertwined with the quality and accessibility of underlying lexical representations. In dyslexia, models emphasize the crucial role of high-quality lexical representations (Perfetti, 2007, 2019), such that word recognition difficulties in dyslexia stem from not only imprecise phonological representation but also lexical integration and heightened lexical competition (Goranova et al., 2021; Jones et al., 2016), as opposed to a global inhibitory impairment. Our results extend previous findings of increased semantically related errors in dyslexic children during picture naming (Nation et al., 2001) to dyslexic adults, highlighting that the pattern of findings observed in the current study is reflective of increased lexical-semantic competition and for a subset of dyslexic individuals, low lexical precision. This impairment could arise due to an over-reliance on lexical and contextual information to compensate for the grapheme-to-phoneme mappings, thus leading to increased lexical activation and competition.

Similarly, our results for AWS are novel, the standard Stroop has not been used previously applied to this specific population. Our results of AWS with low lexical precision show higher Stroop interference than neurotypical peers, while those with high lexical precision show comparable Stroop interference as neurotypical peers. The present findings provide insight into a previously observed discrepancy within the literature. For instance, AWS have, in some studies, been found to take longer to name pictures than controls (e.g., Maxfield, 2015), but other studies have reported no group differences in picture naming tasks (e.g., Hennessey et al., 2014; Maxfield, 2020, 2021; Newman & Bernstein Ratner, 2007). This suggests that this inconsistency may be explained by variations in lexical precision, which could exacerbate disfluencies. The process of accessing a specific word requires a speaker to select a single lexical unit from a pool of co-activated alternatives. In AWS with lower lexical precision, this competition is intensified because multiple words are activated more broadly and with less differentiation. This increased internal competition during speech planning could be a source of the observed disfluencies. This perspective aligns with competitive models of lexical selection, such as WEAVER++ (Levelt et al., 1999; Roelofs, 2018), where a lexical unit is selected only when its activation level surpasses that of its competitors (Roelofs, 2018).

From a practical standpoint, our findings argue for a more nuanced approach to assessment and intervention, particularly in clinical populations. While the Stroop task is a well-established measure of inhibitory control, our results suggest that higher Stroop interference may not always reflect a generalized executive function impairment. Instead, the specific contribution of lexical precision—the quality of an individual's lexical-semantic representations—should be considered (e.g., Protopapas et al., 2007). This finding has important implications for both research and clinical practice. Researchers should incorporate measures that reflect the lexical precision construct (e.g., print exposure, spelling) to better understand how it moderates Stroop interference. By doing so, they can move beyond attributing all Stroop performance to a singular measure of inhibitory control. Similarly, researchers should be mindful that a person's performance on the Stroop task may reflect underlying differences in lexical-semantic processing, especially in populations where these processes are known to be impaired, such as those with dyslexia or AWS. By incorporating measures that reflect lexical precision alongside traditional cognitive assessments, researchers can gain a more comprehensive understanding of an individual’s cognitive profile (Viesel-Nordmeyer et al., 2023). This multi-faceted approach allows for more personalized and targeted reading or speech fluency interventions. For instance, rather than focusing solely on executive control, such interventions could explicitly target the development and strengthening of high-quality lexical representations. This would appear to be a crucial step in improving the efficacy of such interventions, as these representations appear to directly improve an individual's ability to resolve the lexical conflict inherent in the Stroop task and, by extension, in everyday reading and speech.

Limitations

The current study has some limitations that should be addressed in future investigations. First, our analyses did not reveal an association between lexical precision and inhibitory control in neurotypical adults when examining groups individually. We note, however, that the statistical power for these correlational analyses was limited by the sample size within each of the three groups. Therefore, larger sample sizes will be crucial in future investigations to improve the generalizability of our current findings. While we acknowledge the challenges inherent in assembling a substantial cohort of adults who stutter, a group particularly difficult to recruit for research, we recognize that expanding our sample sizes will be crucial for validating and extending our current conclusions. Nevertheless, when we analyzed the entire sample, the relationship between lexical precision and Stroop interference showed no association with conflict resolution. Based on this, we tentatively conclude that lexical precision does not appear to be related to conflict resolution in neurotypical adults during the

Stroop task. In contrast, this relationship *does* seem to be present in AWS and dyslexic adults. Considering previous research findings alongside the present study's sensitivity limitations, our results underscore the need for further and larger scale research to develop a more comprehensive understanding of the relationship between conflict resolution and lexical precision in neurotypical adults, AWS, and dyslexic adults. Second, the Stroop task only assesses one aspect of executive function, inhibition, and not the other core areas of executive function such as updating and set-shifting (e.g., Diamond, 2013; Miyake et al., 2000). As a result, we suggest that future researchers use a suite of individual differences measures from the present study to assess executive function measures to ascertain which core components of executive function are impaired in both groups.

In conclusion, the current study explored the role of inhibitory control in dyslexic adults and AWS, providing novel insights into executive functions within these populations. Previous research has largely focused on distinct aspects of executive function but the current investigation specifically examined the role of lexical precision in inhibitory control within these neurodevelopmental profiles. We observed that dyslexic adults and AWS showed a larger Stroop interference than neurotypical adults. This Stroop interference was modulated by the component of lexical precision: individuals demonstrating high lexical precision showed Stroop interference comparable to neurotypical readers, whereas those with lower spelling proficiency and reduced print exposure showed greater interference. Collectively, these findings underscore the intricate relationship between lexical skills, reading experience and inhibitory control in these populations.

Acknowledgements

We gratefully acknowledge the Stammer Trust for their financial support for participant reimbursement and crucial assistance with the recruitment of adults who stutter. We also thank STAMMA and Action for Stammering in Children (ASC) for their help in recruiting adults who stutter. Finally, we are grateful to the Accessibility Teams at the Universities of

Leicester, Birmingham, and Warwick for their help with the recruitment of dyslexic adults. We especially thank all the participants for their valuable contributions to this work.

Declarations of Interest: None

Data availability statement: The materials and the data sets generated and analysed are available on the Open Science Framework (OSF) repository, https://osf.io/mbc43. This repository also includes an R Markdown script to reproduce all analyses and generate the manuscript.

CReDiT authorship contribution statement

Mahmoud Elsherif: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Jamie Smith-Spark**: Project administration, Supervision, Visualisation, Writing review & editing. **Linda Wheeldon**: Conceptualisation, Project Administration, Supervision, Writing - review and editing. **Steven Frisson**: Conceptualisation, Project Administration, Supervision, Writing - review and editing.

Funding Statement: The Stammer Trust provided financial support to MME for the recruitment of adults who stutter.

References

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behaviour Research Methods, 40*(1), 278289. https://doi.org/10.3758/BRM.40.1.278

Algom, D., & Chajut, E. (2019). Reclaiming the Stroop effect back from control to input-driven attention and perception. *Frontiers in psychology, 10*, 1683. https://doi.org/10.3389/fpsyg.2019.01683

Anderson, J. D., & Wagovich, S. A. (2010). Relationships among linguistic processing speed, phonological working memory, and attention in children who stutter. *Journal of fluency disorders*, *35*(3), 216-234. https://doi.org/10.1016/j.jfludis.2010.04.003

Anderson, J. D., & Wagovich, S. A. (2017). Explicit and implicit verbal response inhibition in preschool-age children who stutter. *Journal of Speech, Language, and Hearing Research, 60*(4), 836-852. https://doi.org/10.1044/2016_JSLHR-S-16-0135

Andrews, S., & Bond, R. (2009). Lexical expertise and reading skill: Bottom-up and top-down processing of lexical ambiguity. *Reading and Writing, 22*(6), 687-711. https://doi.org/10.1007/s11145-008-9137-7

Andrews, S. (2012). Individual differences in skilled visual word recognition: The role of lexical quality. In J. S. Adelman (Ed.), *Visual word recognition* (Vol. 2, pp. 151–172). Psychology Press.

Andrews, S. (2015). Individual differences among skilled readers: The role of lexical quality. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 129–148). Oxford University Press.

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General, 139*(2), 299–318. https://doi.org/10.1037/a0018366

Andrews, S., & Lo, S. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 152–163. https://doi.org/10.1037/ a0024953

Andrews, S., & Lo, S. (2013). Is morphological priming stronger for transparent than opaque words? It depends on individual differences in spelling and vocabulary. *Journal of Memory and Language, 68*(3), 279–296. https://doi.org/10.1016/j. Jml.2012.12.001

Anderson, J. D., & Ofoe, L. C. (2019). The role of executive function in developmental stuttering. *Seminars in Speech and Language, 40*(4), 305–319. https://doi.org/10.1055/s-0039-1692965

Bahrami, H., Nejati, V., & Pooretemad, H. (2014). A comparative study of phonemic and semantic verbal fluency in children and adolescents with developmental stuttering. *Zahedan Journal of Research in Medical Sciences, 16*, 41–44.

Barr D. J., Levy R., Scheepers C., Tily H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates D., Maechler M., Dai B. (2010). *lme4: Linear mixed-effects models using S4 classes* (R package Version 0.999375-27) [Computer software]. R Core Team.

Bednarek, D. B., Saldaña, D., Quintero-Gallego, E., García, I., Grabowska, A., & Gómez, C. M. (2004). Attentional deficit in dyslexia: a general or specific impairment?. *Neuroreport, 15*(11), 1787-1790. https://doi.org/10.1097/01.wnr.0000134843.33260.bf

Bernaards C. A., Jennrich R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement, 65*(5), 676–696. https://doi.org/10.1177/0013164404272507

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and individual differences, 21*(4), 327-336. https://doi.org/10.1016/j.lindif.2011.01.007

Booth, J. N., Boyle, J. M., & Kelly, S. W. (2010). Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: findings from a meta-analysis. *British Journal of Developmental Psychology, 28*(1), 133–176. https://doi.org/10.1348/026151009x485432

Booth, J. N., Boyle, J. M., & Kelly, S. W. (2014). The relationship between inhibition and working memory in predicting children's reading difficulties. *Journal of research in reading, 37*(1), 84-101. https://doi.org/10.1111/1467-9817.12011

Bowers, A., Bowers, L. M., Hudock, D., & Ramsdell-Hudock, H. L. (2018). Phonological working memory in developmental stuttering: potential insights from the neurobiology of language and cognition. *Journal of fluency disorders, 58*, 94-117. https://doi.org/10.1016/j.jfludis.2018.08.006

Bowey, J. A., & Rutherford, J. (2007). Imbalanced word-reading profiles in eighth-graders. *Journal of Experimental Child Psychology, 96*(3), 169-196. https://doi.org/10.1016/j.jecp.2006.11.001

Brosnan, M., Demetre, J., Hamill, S., Robson, K., Shepherd, H., & Cody, G. (2002). Executive functioning in adults and children with developmental dyslexia. *Neuropsychologia, 40,* 2144-2155. https://doi.org/10.1016/s0028-3932(02)00046-5

Brown, T. E., & Landgraf, J. M. (2010). Improvements in Executive Function Correlate with Enhanced Performance and Functioning and Health-Related Quality of Life: Evidence from 2 Large, Double-Blind, Randomized, Placebo-Controlled Trials in ADHD. *Postgraduate Medicine, 122*(5), 42–51. https://doi.org/10.3810/pgm.2010.09.2200

Burca, M., Beaucousin, V., Chausse, P., Ferrand, L., Parris, B. A., & Augustinova, M. (2021). Is there semantic conflict in the Stroop task?: Further evidence from a modified twoto-one Stroop paradigm combined with single-letter coloring and cueing. *Experimental Psychology, 68*(5), 274–283. https://doi.org/10.1027/1618-3169/a000530

Burca, M., Chausse, P., Ferrand, L., Parris, B. A., & Augustinova, M. (2022). Some further clarifications on age-related differences in the Stroop task: New evidence from the two-to-one Stroop paradigm. *Psychonomic Bulletin & Review, 29*(2), 492–500. https://doi.org/10.3758/s13423-021-02011-x

Burt, J. S., & Fury, M. B. (2000). Spelling in adults: The role of reading skills and experience. *Reading and Writing, 13*(12), 1–30. https://doi.org/10.1023/A:1008071802996

Castles, A., & Friedmann, N. (2014). Developmental Dyslexia and the Phonological Deficit Hypothesis. *Mind and Language, 29,* 270-285. https://doi.org/10.1111/mila.12050

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & cognition, 28*(1), 143-153. https://doi.org/10.3758/BF03211582

Cohen J. (1988). *Statistical power analysis for the behavioral sciences.* Erlbaum-Drucker.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological review, 97*(3), 332. https://doi.org/10.1037/0033-295X.97.3.332

Coltheart, M., Woollams, A., Kinoshita, S., & Perry, C. (1999). A position-sensitive Stroop effect: Further evidence for a left-to-right component in print-to-speech conversion. *Psychonomic Bulletin & Review, 6*(3), 456-463. https://doi.org/10.3758/BF03210835

Cortese, M. J., & Simpson, G. B. (2000). Regularity effects in word naming: What are they?. *Memory & Cognition, 28*(8), 1269-1276. https://doi.org/10.3758/BF03211827

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology, 82*(4), 733-740. https://doi.org/10.1037/0022-0663.82.4.733

Dalrymple-Alford, E. C. (1972). Associative facilitation and interference in the Stroop color-word task. *Perception & Psychophysics, 11*(4), 274-276. https://doi.org/10.3758/BF03210377

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Demoulin, C., & Kolinsky, R. (2016). Does learning to read shape verbal working memory?. *Psychonomic bulletin & review, 23*(3), 703-722. https://doi.org/10.3758/s13423-015-0956-7

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64,* 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dobó, D., Ladányi, E., Szőllősi, Á., Lukics, K. S., Németh, K., & Lukács, Á. (2022). The relationship between cognitive control and lexical conflict resolution in developmental dyslexia. *Clinical Linguistics & Phonetics, 3*6(4-5), 319–340. https://doi.org/10.1080/02699206.2021.1998632

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British picture vocabulary scale* (2nd ed.). Windsor, UK: NFER-Nelson.

Eggers, K., De Nil, L. F., & Van den Bergh, B. R. (2012). The efficiency of attentional networks in children who stutter. *Journal of Speech, Language, and Hearing Research, 55,* 946–959. https://doi.org/10.1044/1092-4388(2011/10-0208)

Eggers, K., De Nil, L. F., & Van den Bergh, B. R. (2013). Inhibitory control in childhood stuttering. *Journal of Fluency Disorders, 38*, 1–13. https://doi.org/10.1016/j.jfludis.2012.10.001

Eichorn, N., Marton, K., & Pirutinsky, S. (2018). Cognitive flexibility in preschool children with and without stuttering disorders. *Journal of Fluency Disorders, 57*, 37-50. https://doi.org/10.1016/j.jfludis.2017.11.001

Elliott, C. D., Smith, P., & McCulloch, K. (1996). *British ability scales second edition* (BAS II). Windsor, Berkshire: NFER-Nelson.

Elsherif, M. M., Frisson, S., & Wheeldon, L. R. (2023). Orthographic precision for word naming in skilled readers. *Language, Cognition and Neuroscience, 38*(2), 197-216. https://doi.org/10.1080/23273798.2022.2108091

Elsherif, M. M., Wheeldon, L. R., & Frisson, S. (2021). Do dyslexia and stuttering share a processing deficit? *Journal of Fluency Disorders, 67*, Article 105827. https://doi.org/10.1016/j. jfludis.2020.105827

Elsherif, M. M., Wheeldon, L. R., & Frisson, S. (2022). Phonological precision for word recognition in skilled readers. *Quarterly Journal of Experimental Psychology, 75*(6), 10211040. https://doi.org/10.1177/17470218211046350

Everatt, J., Warner, J., Miles, T. R., & Thomson, M. E. (1997). The incidence of Stroop interference in dyslexia. *Dyslexia, 3*(4), 222–228. https://doi.org/10.1002/(sici)1099-0909(199712)3:4%3C222::aid-dys12%3E3.0.co;2-p

Facoetti, A., & Turatto, M. (2000). Asymmetrical visual fields distribution of attention in dyslexic children: a neuropsychological study. *Neuroscience letters, 290*(3), 216-218. https://doi.org/10.1016/S0304-3940(00)01354-9

Faccioli, C., Peru, A., Rubini, E., & Tassinari, G. (2008). Poor readers but compelled to read: Stroop effects in developmental dyslexia. *Child Neuropsychology, 14*(3), 277-283. https://doi.org/10.1080/09297040701290040

Field A. P. (2009). *Discovering statistics using SPSS: (And sex and drugs and rock "n" roll)* (3rd ed.). SAGE.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Goranova, Z. O., Olson, A. C., & Krott, A. (2021). What about lexical competition? Exploring the locus of lexical retrieval deficits in adults with developmental dyslexia. *Neuropsychology, 35*(8), 822–846. https://doi.org/10.1037/neu0000767

Guitar, B. (2013). *Stuttering: An integrated approach to its nature and treatment.* Lippincott Williams & Wilkins.

Hall, J., McGregor, K. K., & Oleson, J. (2017). Weaknesses in lexical-semantic knowledge among college students with specific learning disabilities: Evidence from a semantic fluency task. *Journal of Speech, Language, and Hearing Research, 60*(3), 640-653. https://doi.org/10.1044/2016_JSLHR-L-15-0440

Handler, S. M., Fierson, W. M., Section on Ophthalmology, Council on Children with Disabilities, American Academy of Ophthalmology, American Association for Pediatric Ophthalmology and Strabismus, & American Association of Certified Orthoptists (2011). Learning disabilities, dyslexia, and vision. *Pediatrics, 127*(3), e818–e856. https://doi.org/10.1542/peds.2010-3670

Hasshim, N., & Parris, B. A. (2021). The role of contingency and correlation in the Stroop task. *Quarterly Journal of Experimental Psychology, 74(*10), 1657-1668.https://doi.org/10.1177/17470218211032548

Helland, T., & Asbjørnsen, A. (2000). Executive functions in dyslexia. *Child Neuropsychology, 6*(1), 37-48. https://doi.org/10.1076/0929-7049(200003)6:1;1-B;FT037

Henderson, S. K., Peterson, K. A., Patterson, K., Lambon, M. A., & Rowe, J. B. (2023). Verbal fluency tests assess global cognitive status but have limited diagnostic differentiation: evidence from a large-scale examination of six neurodegenerative diseases. *Brain Communications, 5*(2). https://doi.org/10.1093/braincomms/fcad042

Hennessey, N. W., Dourado, E., & Beilby, J. M. (2014). Anxiety and speaking in people who stutter: An investigation using the emotional Stroop task. *Journal of fluency disorders, 40*, 44-57. https://doi.org/10.1016/j.jfludis.2013.11.001

Holmes J., Hilton K. A., Place M., Alloway T. P., Elliott J. G., Gathercole S. E. (2014). Children with low working memory and children with ADHD: Same or different? *Frontiers in Human Neuroscience, 8*, Article 976. https://doi.org/10.3389/fnhum.2014.00976

Howell, P., & Au-Yeung, J. (2002). The EXPLAIN theory of fluency control applied to the diagnosis of stuttering. In E. Fava (Ed.), *Clinical linguistics: Theory and applications in speech pathology and therapy* (Vol. 4, pp. 75–94). John Benjamins. https://doi.org/10.1075/cilt.227.08how

Huettig, F., Lachmann, T., Reis, A., & Petersson, K. M. (2018). Distinguishing cause from effect–many deficits associated with developmental dyslexia may be a consequence of reduced and suboptimal reading experience. *Language, Cognition and Neuroscience, 33*(3), 333-350. https://doi.org/10.1080/23273798.2017.1348528

Huettig, F., & Pickering, M. J. (2019). Literacy advantages beyond reading: Prediction of spoken language. *Trends in Cognitive Sciences, 23*(6), 464–475. https://doi.org/10.1016/j.tics.2019.03.008

Hulme, C., & Snowling, M. J. (2009). *Developmental disorders of language learning and cognition*. Wiley Blackwell.

Iles, J., Walsh, V., & Richardson, A. (2000). Visual search performance in dyslexia. *Dyslexia, 6*(3), 163-177. https://doi.org/10.1002/1099-0909(200007/09)6:3%3C163::AID-DYS150%3E3.0.CO;2-

Jeffries, S., & Everatt, J. (2004). Working memory: its role in dyslexia and other specific learning difficulties. *Dyslexia*, *10*(3), 196–214. https://doi.org/10.1002/dys.278

Jones, M. W., Branigan, H. P., Hatzidaki, A., & Obregón, M. (2010). Is the ‘naming’ deficit in dyslexia a misnomer?. *Cognition, 116*(1), 56-70. https://doi.org/10.1016/j.cognition.2010.03.015

Jones, M. W., Snowling, M. J., & Moll, K. (2016). What automaticity deficit? Activation of lexical information by readers with dyslexia in a rapid automatized naming Stroop-switch task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(3), 465–474. https://doi.org/10.1037/xlm0000186

Kakuta, K., & Kawasaki, A. (2025). Executive function in preschool children who stutter: a behavioral assessment study. *Scientific Reports, 15*(1), 16159. https://doi.org/10.1038/s41598-025-00981-9

Kapoula, Z., Lê, T. T., Bonnet, A., Bourtoire, P., Demule, E., Fauvel, C., Quilicci, C., & Yang, Q. (2010). Poor Stroop performances in 15-year-old dyslexic teenagers. *Experimental Brain Research, 203*(2), 419–425. https://doi.org/10.1007/s00221-010-2247-x

Kass, R. A., & Tinsley, H. E. A. (1979). Factor analysis. *Journal of Leisure Research, 11,* 120–138.

Khan A. (2014). An investigation into prospective memory in children with developmental dyslexia. *Frontiers in Psychology, 5,* 1308. https://doi.org/10.3389/fpsyg.2014.01308

Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *The American journal of psychology, 77*(4), 576-588. https://doi.org/10.2307/1420768

Laberge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293–323. https://doi.org/10.1016/00100285(74)90015-2

Lescht, E., Dickey, M. W., Stockbridge, M. D., & Ratner, N. B. (2022). Adults who stutter show diminished word fluency, regardless of mode. *Journal of Speech, Language, and Hearing Research, 65*(3), 906-922. https://doi.org/10.1044/2021_JSLHR-21-00344

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences, 22*(1), 1-38. https://doi.org/10.1017/S0140525X99001776

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly, 13*(2), 123–146. https://doi.org/10.1080/1057356970130203

Lonergan, A., Doyle, C., Cassidy, C., MacSweeney Mahon, S., Roche, R. A. P., Boran, L., & Bramham, J. (2019). A meta-analysis of executive functioning in dyslexia with consideration of the impact of comorbid ADHD. *Journal of Cognitive Psychology, 31*(7), 725–749. https://doi.org/10.1080/20445911.2019.1669609

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163

Mahé, G., Doignon-Camus, N., Dufour, A., & Bonnefond, A. (2014). Conflict control processing in adults with developmental dyslexia: an event related potentials study. *Clinical Neurophysiology, 125*(1), 69-76. https://doi.org/10.1016/j.clinph.2013.06.005

Mano, Q. R., Williamson, B. J., Pae, H. K., & Osmon, D. C. (2016). Stroop interference associated with efficient reading fluency and prelexical orthographic processing. *Journal of clinical and experimental neuropsychology, 38*(3), 275-283. https://doi.org/10.1080/13803395.2015.1107029

Martin-Chang, S. L., & Gould, O. N. (2008). Revisiting print exposure: Exploring differential links to vocabulary, comprehension and reading rate. *Journal of Research in Reading, 31*(3), 273–284. https://doi.org/10.1111/j.14679817.2008.00371.x

Maxfield, N. D. (2020). Inhibitory control of lexical selection in adults who stutter. *Journal of Fluency Disorders, 66*, 105780. https://doi.org/10.1016/j.jfludis.2020.105780

Maxfield, N. D. (2021). Cognitive control of action naming in adults who stutter. *Journal of fluency disorders, 70*, 105841. https://doi.org/10.1016/j.jfludis.2021.105841

Maxfield, N. D., Morris, K., Frisch, S. A., Morphew, K., & Constantine, J. L. (2015). Real-time processing in picture naming in adults who stutter: ERP evidence. *Clinical Neurophysiology, 126*(2), 284-296. https://doi.org/10.1016/j.clinph.2014.05.009

McKenna, F. P., & Sharma, D. (1995). Intrusive cognitions: An investigation of the emotional Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(6), 1595–1607. https://doi.org/10.1037/0278-7393.21.6.1595

McLean, G. M., Stuart, G. W., Coltheart, V., & Castles, A. (2011). Visual temporal processing in dyslexia and the magnocellular deficit theory: the need for speed?. *Journal of Experimental Psychology: Human Perception and Performance, 37*(6), 1957. https://doi.org/10.1037/a0024668

Megherbi, H., Elbro, C., Oakhill, J., Segui, J., & New, B. (2018). The emergence of automaticity in reading: Effects of orthographic depth and word decoding ability on an adjusted Stroop measure. *Journal of experimental child psychology, 166*, 652-663. https://doi.org/10.1016/j.jecp.2017.09.016

Menghini, D., Finzi, A., Carlesimo, G. A., & Vicari, S. (2011). Working memory impairment in children with developmental dyslexia: is it just a phonological deficit?. *Developmental Neuropsychology, 36*(2), 199–213. https://doi.org/10.1080/87565641.2010.549868

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager T. D. (2000). The unity and diversity of executive functions, and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49-100. https://doi.org/10.1006/cogp.1999.0734

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition, 24*(1-2), 45-64. https://doi.org/10.1016/0010-0277(86)90004-1

Monsell, S., Taylor, T. J., & Murphy, K. (2001). Naming the color of a word: Is it responses or task sets that compete?. *Memory & Cognition*, *29*(1), 137-151. https://doi.org/10.3758/BF03195748

Nation, K., Marshall, C. M., & Snowling, M. J. (2001). Phonological and semantic contributions to children's picture naming skill: Evidence from children with developmental reading disorders. *Language and cognitive processes, 16*(2-3), 241-259. https://doi.org/10.1080/01690960042000003

Newman, R. S., & Ratner, N. B. (2007). The role of selected lexical factors on confrontation naming accuracy, speed, and fluency in adults who do and do not stutter. *Journal of Speech, Language, and Hearing Research, 50*(1), 196-213. https://doi.org/10.1044/1092-4388(2007/016)

Nicolson, R. I., & Fawcett, A. J. (1990). Automaticity: A new framework for dyslexia research?. *Cognition, 35*(2), 159-182. https://doi.org/10.1016/0010-0277(90)90013-A

Ning, N., Peng, D., Liu, X., & Yang, S. (2017). Speech timing deficit of stuttering: evidence from contingent negative variations. *PloS one, 12*(1), e0168836. https://doi.org/10.1371/journal.pone.0168836

O'Brien, G., & Yeatman, J. D. (2021). Bridging sensory and language theories of dyslexia: Toward a multifactorial model. *Developmental Science, 24*(3), e13039. https://doi.org/10.1111/desc.13039

Ofoe, L. C., Anderson, J. D., & Ntourou, K. (2018). Short-term memory, inhibition, and attention in developmental stuttering: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 61*(7), 1626-1648. https://doi.org/10.1044/2018_JSLHR-S-17-0372

Ofoe, L. C., Katerina Ntourou, Clifton, S., & Coalson, G. A. (2025). A Meta-Analytical Review of Executive Function Skills in Adults who Stutter. *MedRxiv*. https://doi.org/10.1101/2025.09.02.25334917

Parris, B. A., Augustinova, M., & Ferrand, L. (2019). The locus of the stroop effect. *Frontiers in Psychology, 10*, 2860. https://doi.org/10.3389/fpsyg.2019.02860

Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (2022). The loci of Stroop effects: a critical review of methods and evidence for levels of processing contributing to color-word Stroop effects and the implications for the loci of attentional selection. *Psychological Research*, *86*(4), 1029–1053. https://doi.org/10.1007/s00426-021-01554-x

Pearson, A. D., Grant, A., Leigh, J., Macdonald, S. J., Williams, K., Williams, G. L., Spaeth, E., Rose, K., Memmott, A., & Botha, M. (2025, August 1). “It’s like calling short people vertically challenged”: Language and terminology preferences among neurodivergent adults. *PsyArXiv*. https://doi.org/10.31219/osf.io/jf62e_v1

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Lawrence Erlbaum.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. Scientific Studies of Reading, 11(4), 357–383. https://doi.org/10.1080/10888430701530730

Perfetti, C. A. (2017). Lexical quality revisited. In E. Segers & P. van den Broek (Eds.), *Developmental perspectives in written language and literacy: In honor of Ludo Verhoeven* (pp. 51–68). John Benjamins.

Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. S. Gorfien (Ed.), *On the consequences of meaning selection (*pp. 67–86). American Psychological Association.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven, C. Elbron, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). John Benjamins.

Perkins, W. H., Kent, R. D., & Curlee, R. F. (1991). A theory of neuropsycholinguistic function in stuttering. *Journal of Speech, Language, and Hearing Research, 34*(4), 734-752. https://doi.org/10.1044/jshr.3404.734

Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *The lancet, 379*(9830), 1997-2007. https://doi.org/10.1016/S0140-6736(12)60198-6

Protopapas, A., Archonti, A., & Skaloumbakas, C. (2007). Reading ability is negatively related to Stroop interference. *Cognitive Psychology, 54*(3), 251-282. https://doi.org/10.1016/j.cogpsych.2006.07.003

Proulx, M. J., & Elmasry, H. M. (2015). Stroop interference in adults with dyslexia. *Neurocase*, *21*(4), 413-417. https://doi.org/10.1080/13554794.2014.914544

Provazza, S., Adams, A.-M., Giofrè, D., & Roberts, D. J. (2019). Double trouble – visual and phonological impairments in English dyslexic readers. *Frontiers in Psychology, 10,* Article 2725. https://doi.org/10.3389/fpsyg.2019.02725

R Development Core Team. (2017). *r: A language and environment for statistical computing.* R Foundation for Statistical Computing. http://www.R-project.org

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain, 126*(Pt 4), 841–865. https://doi.org/10.1093/brain/awg076

Raven J. C. (1960). *Guide to using the standard progressive matrices.* Lewis.

Reiter, A., Tucha, O., & Lange, K. W. (2005). Executive functions in children with dyslexia. *Dyslexia, 11*(2), 116-131. https://doi.org/10.1002/dys.289

Reynolds, C. R., & Kamphaus, R. W. (2007). *TIWRE: Test of irregular word reading efficiency.* Psychological Assessment Resources.

Roelofs, A. (2018). A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition, 172*, 59-72. https://doi.org/10.1016/j.cognition.2017.12.007

Scarpina, F., & Tagini, S. (2017). The Stroop Color and Word Test. *Frontiers in Psychology*, *8*, 557. https://doi.org/10.3389/fpsyg.2017.00557

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*(2), 151-218. https://doi.org/10.1016/0010-0277(94)00645-2

Smith-Spark, J. H., & Fisk, J. E. (2007). Working memory functioning in developmental dyslexia. *Memory, 15,* 34-56. https://doi.org/10.1080/09658210601043384

Smith-Spark, J. H., & Gordon, R. (2022). Automaticity and executive abilities in developmental dyslexia: A theoretical review. *Brain Sciences, 12,* 446. https://doi.org/10.3390/brainsci12040446

Smith-Spark, J. H., Henry, L. A., Messer, D. J., Edvardsdottir, E., & Zięcik, A. P. (2016). Executive functions in adults with developmental dyslexia. *Research in Developmental Disabilities, 53-54,* 323-341. https://doi.org/10.1016/j.ridd.2016.03.001

Smith-Spark, J. H., Henry, L. A., Messer, D. J., & Zięcik, A. P. (2017). Verbal and non-verbal fluency in adults with developmental dyslexia: Phonological processing or executive control problems? *Dyslexia, 23,* 234-250. https://doi.org/10.1002/dys.1558

Smith-Spark, J. H., & Moore, V. (2009). The representation and processing of familiar faces in dyslexia: Differences in age of acquisition effects. *Dyslexia, 15*(2), 129-146. https://doi.org/10.1002/dys.365

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging, 14*(3), 458–472. https://doi.org/10.1037/0882-7974.14.3.458

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading research quarterly,* 402-433. https://doi.org/10.2307/747605

Stein, J. (2023). Theories about Developmental Dyslexia. *Brain Sciences, 13*(2), 208. https://doi.org/10.3390/brainsci13020208

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662. https://doi.org/10.1037/h0054651

Stubenrauch, C., Freund, J., Alecu de Flers, S., Scharke, W., Braun, M., Jacobs, A. M., & Konrad, K. (2014). Nonword reading and Stroop interference: What differentiates attention-deficit/hyperactivity disorder and reading disability?. *Journal of clinical and experimental neuropsychology, 36*(3), 244-260. https://doi.org/10.1080/13803395.2013.878690

Tabachnick B. G., Fidell L. S., Ullman J. B. (2007). *Using multivariate statistics* (Vol. 5). Pearson.

Treleaven, S. B., & Coalson, G. A. (2020). Manual response inhibition and quality of life in adults who stutter. *Journal of Communication Disorders, 88*, Article 106053. https://doi.org/10.1016/j.jcomdis.2020.106053

Torgesen, J. K., Rashotte, C. A., & Wagner, R. K. (1999). *TOWRE: Test of word reading efficiency*. Austin, TX: Pro-ed.

Uttl, B., & Graf, P. (1997). Color-Word Stroop test performance across the adult life span. *Journal of Clinical and Experimental Neuropsychology*, *19*(3), 405–420. https://doi.org/10.1080/01688639708403869

van Lieshout, P., Ben-David, B., Lipski, M., & Namasivayam, A. (2014). The impact of threat and cognitive stress on speech motor control in people who stutter. *Journal of fluency disorders, 40*, 93-109. https://doi.org/10.1016/j.jfludis.2014.02.003

Varvara, P., Varuzza, C., Sorrentino, A. C., Vicari, S., & Menghini, D. (2014). Executive functions in developmental dyslexia. *Frontiers in human neuroscience, 8*, 120. https://doi.org/10.3389/fnhum.2014.0012

Veldre, A., & Andrews, S. (2014). Lexical quality and eye movements: Individual differences in the perceptual span of skilled adult readers. *The Quarterly Journal of Experimental Psychology, 67*(4), 703–727. https://doi.org/10.1080/17470 218.2013.826258

Vellutino, F. R. (1979). *Dyslexia: Theory and research*. Cambridge, MA: MIT Press.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45,* 2-40. https://doi.org/10.1046/j.0021-9630.2003.00305.x

Viesel-Nordmeyer, N., Reuber, J., Kuhn, J. T., Moll, K., Holling, H., & Dobel, C. (2023). Cognitive profiles of children with isolated and comorbid learning difficulties in reading and math: A meta-analysis. *Educational Psychology Review, 35*(1), 34. https://doi.org/10.1007/s10648-023-09735-3

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (1999). *Comprehensive test of phonological processing: CTOPP.* Austin, TX: Pro-ed.

Westfall J., Kenny D. A., Judd C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of*

Experimental Psychology: General, 143, 2020–2045. https://doi.org/10.1037/xge0000014

Wiederholt, J. L., & Blalock, G. (2000). *GSRT: Gray Silent Reading Tests.* Pro-Ed..

Williams, K. T., & Williams, K. T. (2007). *EVT-2: Expressive vocabulary test.* Pearson Assessment.

Winter, B. (2019). *Statistics for linguists: An introduction using R*. Routledge.

Yairi, E., & Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of fluency disorders, 38*(2), 66-87. https://doi.org/10.1016/j.jfludis.2012.11.002