Original Article

QJEP

Quarterly Journal of Experimental Psychology
1–25

DOI: 10.1177/17470218241310440
qjep.sagepub.com
Sage

Effects of central vs. peripheral attentional-oculomotor exercise on lexical processing

Shaylyn Kress[1], Scott Caron[1], Josh Neudorf[2], Braedyn Borowsky[1] and Ron Borowsky[1]

Abstract

Past research from our lab has suggested visual demands in video games serve to exercise attentional-oculomotor (A-O) processing in a manner beneficial to reading. However, testing the effect of video games on reading typically requires long timeframes (e.g., multiweek training or years of accumulated video game experience). The current study manipulated within-experiment peripheral and central demands to evaluate the effects of A-O exercise on task performance. Our study included two tasks: an orthographic lexical decision task (OLDT), designed to optimise orthographic lexical processing, and a novel graphic-based health bar decision task (HBDT). In Experiment 1, the stimuli were presented centrally in one block and peripherally in another block to manipulate A-O exercise. We observed greater improvements in the peripheral-first than the central-first group, particularly for the OLDT. In Experiments 2 and 3, we focused on the OLDT, with the HBDT serving as the A-O exercise task, and observed improvements in both centrally and peripherally trained participants. We additionally observed, through analyses of word and bigram frequency, a double dissociation, whereby increased target word frequency was associated with faster target reaction times and improved error rates, whereas increased foil bigram frequency was associated with slower foil reaction times and worse error rates. Taken together, the experiments demonstrate a mechanism beyond simple task learning that drives reading improvements, and A-O exercise, even if movements are small, appears to play a role in the improvements observed. We suggest future research should further develop this paradigm and examine its utility for reading remediation in dyslexia.

Keywords

Visual-spatial attention; reading; lexical decision; oculomotor processing; orthographic lexical processing; sublexical processing; word frequency; bigram frequency

Received: 25 July 2024; revised: 4 November 2024; accepted: 6 November 2024

When an athlete wants to improve at activities like weightlifting, or running, they will prepare a training regimen of exercises targeted to strengthen their muscles or their endurance. Nonathletes also engage in physical exercise to maintain their form and health, and just as lifting weights or practicing for a marathon improves your physical fitness, one can also engage in cognitive exercise to maintain or improve mental skills such as memory (Holmes et al., 2019) or arithmetic (Sella et al., 2016). There is also commercial interest in cognitive exercise, with mobile applications and games purporting to improve cognitive skills. Nintendo's game, *BrainAge*, is one example of a commercial game which aims to improve cognitive skills through a variety of mini-games (Nouchi et al., 2012). In the case of *BrainAge*, researchers observed improvements in executive function and processing speed relative to the control group, thus demonstrating the potential for commercial video games to improve cognitive skills.

Reading is an incredibly valuable cognitive skill and uniquely human experience. Printed text is a common occurrence on signage, medication labels, employment contracts, or bank documents, which all need to be read and comprehended to act on appropriately. Difficulties with reading therefore lead to difficulties in carrying out many day-to-day tasks required for success and well-being. The goal of our present research is to test whether

---

[1]University of Saskatchewan, Saskatoon, Saskatchewan, Canada
[2]Simon Fraser University, Burnaby, British Columbia, Canada

Corresponding author:
Ron Borowsky, Department of Psychology and Health Studies, University of Saskatchewan, 9 Campus Dr., Saskatoon, Saskatchewan S7N 5A5, Canada.
Email: ron.borowsky@usask.ca

reading can be improved through cognitive exercise, namely attentional-oculomotor (A-O) exercise.

Reading requires cognitive exercise

Improving literacy is a global issue; the international median rate of students at or below the low reading benchmark was 16% (Brochu et al., 2018). Both children and adults can struggle with literacy. The Council of Ministers of Education Canada and Statistics Canada (2013) reported that 17% of adult Canadians are at the lowest levels for literacy.

Given the vital role that reading plays in everyday activities, and the currently existing gaps in many individuals' reading skills, reading makes an excellent candidate to target through cognitive exercise. Extracurricular reading is perhaps the broadest option for improving reading skills, and higher levels of reading engagement are associated with better reading comprehension and vocabulary (Pfost et al., 2013), but this natural form of reading practice is partially reliant on the individual's baseline ability and interest in reading. If individuals with dyslexia have negative associations with the reading experience, any anxiety or frustration that accompanies the activity could discourage them from extracurricular reading (Stanovich, 1986).

Rather than relying on organic reading, which could contribute to existing equity gaps in reading ability, educational organisations could instead use cognitive exercise that specifically targets reading (reading-focused exercise). Reading-focused exercise could include activities such as phonetic decoding (Foorman et al., 2018), text reading, orthographic rule instruction, multisensory instruction (Hall et al., 2022), and individually tailored reading interventions (Partanen et al., 2019). These exercises are often administered over multiple sessions, under the guidance of an educator or specialist, which can help ensure the individual is making regular progress in their reading improvement. However, the extended timeframes and personnel required in most reading interventions may pose barriers for children from impoverished communities and for adults juggling multiple responsibilities.

Alternatively, cognitive exercise could target important processes that make reading easier. Given the multimodal system of reading, this approach opens numerous avenues for activities to indirectly improve reading, particularly through visual attention, which we have already seen is a key component in reading processes (e.g., Ekstrand et al., 2016; Ekstrand, Neudorf, Gould, et al., 2019; Ekstrand, Neudorf, Kress, & Borowsky, 2019; Gabrieli & Norton, 2012).

Visual demands and reading performance

Some previous research from our lab found that greater experience with peripheral demands in video games was associated with faster reading speeds, which may indicate that peripheral visual demands exercise oculomotor control ability (Kress et al., 2023). This same study observed greater experience with central demands in video games was associated with slower reading speeds, demonstrating that visual-spatial demand location exhibits a double dissociation with reading speed.

This double dissociation suggests that visual-spatial demands employ oculomotor control processes differently depending on location, with peripheral demands exercising oculomotor control processes more than central demands. Other researchers (Minissi et al., 2024; West et al., 2013) have also argued that video games can affect oculomotor performance as well. From our findings (Kress et al., 2023), we argued that games with a high quantity of peripheral visual demands (e.g., enemies popping up in unexpected locations on the outskirts of the screen) encourage the player to regularly and systematically move their eyes across all parts of the screen, exercising the full range of oculomotor function. In contrast, when the player only needs to focus on the centre of the screen, their oculomotor function is not fully engaged. If the oculomotor system is not used to being engaged, it will be less prepared for the systematic eye movements required during reading. Based on that previous work, we propose A-O exercise as a form of cognitive exercise which may benefit reading performance through improved oculomotor control.

A-O exercise for reading

The form of cognitive exercise that we are interested in for the purposes of our research is A-O exercise, which we will define as activities that encourage processing in the visual-attentional system and promote oculomotor control. The visual attention system integrates a number of visual and attentional processes, including low-level visual processing, attentional alerting and orienting mechanisms, and higher-level executive control (Corbetta & Shulman, 2002; Fan et al., 2002). The visual-attentional system is traditionally conceptualised as a two-network system (Corbetta & Shulman, 2002). The ventral attention network handles bottom-up control processes, which are stimulus-driven responses that capture attention automatically (e.g., looking automatically to the source of a sudden sound or movement), and includes the inferior frontal and middle frontal gyri. The temporoparietal junction also plays a role in this network and is thought to act as a switching mechanism (interrupting top-down attention, engaging bottom-up attention) when a salient stimulus arrives in an unattended location. The dorsal attention network handles top-down attentional control processes, which are those goal-driven responses that are capturing attention through voluntary effort (e.g., instructions to attend to a certain stimulus). Regions such as the frontal-eye-field (FEF) and intraparietal sulcus/superior parietal lobule are noteworthy regions of the dorsal attention network.

The FEF is particularly important for its role in oculomotor control. Models of oculomotor control describe the system underlying our eye movements, which involves deciding when and to where eye shifts should take place (Tatler et al., 2017). This control over eye movements is critical in reading, where the reader must systematically shift their eyes across the page. Simulations by Li and Pollatsek (2020) demonstrate the role of eye movements in Chinese word processing, supporting the argument for oculomotor control's involvement in reading.

Researchers in the field of dyslexia frequently propose links between attentional network dysfunction and the reading deficits that are the hallmark of the disorder (Boden & Giaschi, 2007; Boros et al., 2016; Facoetti et al., 2000; Vidyasagar & Pammer, 2010). Oculomotor control difficulties would contribute to difficulties with attentional orienting processes, such as maintaining fixation or moving eyes accurately (Boden & Giaschi, 2007; Morrison, 1984). In the reading network, these attentional orienting difficulties would be most likely to translate to difficulties with visual letter processing (orthographic feature encoding) or phonetic decoding (grapheme-to-phoneme conversion), as these processes are more associated with voluntary attention and oculomotor control than whole-word processing via the orthographic lexicon (Ekstrand, Neudorf, Kress, & Borowsky, 2019), and both orthographic and phonological processes are important to successful reading (Acha et al., 2024).

Video game play as A-O exercise

Based on the literature, video game play may be one activity that promotes A-O exercise. Video games expose players to an assortment of visual stimuli both in central and peripheral areas of the screen. The visual stimuli presented in video games can be thought of as highly valid attentional cues—if a stimulus appears on-screen (such as a quest marker that appears on the player's mini-map, a glowing pop-up that indicates it is possible to interact with an object, or a damage notification that points the player in the direction from which they received the attack), it is almost always important and should be attended to. The high-frequency, high-relevance visual stimuli presented peripherally would encourage substantial use of the oculomotor control system, thus contributing to A-O exercise. Training studies have observed oculomotor control system changes in structure and performance following video game play in older adults (Diarra et al., 2019; West et al., 2013). In a structural brain connectivity study, Kowalczyk et al. (2018) found that the connectivity in visual-spatial regions was greater in real-time strategy players compared with people who do not play games regularly, and behavioural studies have suggested that action video game play may improve reading ability in children (e.g., Bertoni et al., 2021; Franceschini et al., 2017; Pasqualotto et al., 2022; see Puccio et al., 2023 for a meta-analysis).

The research on the relationship between game-based A-O exercise and reading ability has focussed on long timeframes. In training studies, participants receive multiple hours of training over the course of several days or weeks (over 10 hr on average; Puccio et al., 2023), whereas cross-sectional studies typically include variables that reflect months or years of accumulated video game experience (e.g., Kress et al., 2023). The current study seeks to determine whether the benefits of A-O exercise on reading performance can be realised in a brief and efficient timeframe. In addition, much of the past research has focused on the phonological benefits of video game experience (e.g., Bertoni et al., 2021, 2024; Franceschini & Bertoni, 2019; Pasqualotto et al., 2022). This study will extend the field's understanding of the role A-O exercise plays in reading by using a lexical decision task (LDT) designed to isolate orthographic lexical processing from sublexical phonological processing, thus filling the knowledge gap in how A-O exercise may affect orthographic processing.

Measures of lexicality and sublexicality

Word frequency has been used in past reading research as a type of measure to help predict how reliant a stimulus might be on either the lexical or sublexical route (when considering dual-route reading models, Borowsky et al., 2013). Word frequency is a measure of how likely a word is to appear in the given corpus. Generally, researchers find that higher word frequency is associated with faster response times (Borowsky & Besner, 1993; Borowsky et al., 2013; Brysbaert & New, 2009), and greater activation in ventral brain regions (Borowsky et al., 2013), and is thus thought to reflect the level of involvement of the lexical route.

Bigram frequency is another corpus-based word characteristic and is a measure of how frequently a pair of letters appears in the corpus (Balota et al., 2007). There are multiple ways to combine the frequencies of each bigram to calculate a single measure for the word. One option is the sum bigram frequency, where the frequencies of each bigram are simply added together. Another option is the mean bigram frequency, where the sum bigram frequency is divided by the number of bigrams in the word. Bigram frequency is less well studied compared with word frequency, but higher bigram frequency is associated with faster reading response times and has been argued to reflect the speed at which spelling can be mapped onto sound, thus representing the efficiency of the dorsal-sublexical grapheme-to-phoneme processing route (Borowsky et al., 2013). Through reading aloud studies, lexical and sublexical processing has been shown to involve distinct regions of activation in the ventral stream and dorsal stream, respectively (Borowsky et al., 2006; Ekstrand, Neudorf, Kress, & Borowsky, 2019).

Not all models of reading employ a dual-route structure. As Owen and Borowsky (2003) describe, single-route models exist which employ only one nonsemantic route from visual text input to phonological representation. The connectionist models proposed by Seidenberg and colleagues (e.g., Harm & Seidenberg, 1999) are an example of this single-route model conceptualisation. Past research has found that single- vs. dual-route models offer different simulations of reading results depending on the task and stimuli, with some single-route models struggling to predict word frequency interaction effects (see Borowsky & Besner, 2006 for discussion of this issue; see also Besner & Borowsky, 2006; Plaut & Booth, 2006).

In the context of the LDT, most designs focus on word stimuli with nonword foils (see Borowsky & Besner, 1993; Masson & Borowsky, 1998). In this task, the participant must decide whether the stimulus is a word or not. These designs may not be ideal for isolating lexical and sublexical processing, as the word and nonword stimuli are not matched on their non-orthographic lexical features (Kress et al., 2021). For example, using nonspecific word targets (e.g., *gave*) and pronounceable but meaningless nonword foils (e.g., *bave*), a participant could rely on any combination of orthographic (spelling), phonological (sound), and semantic (meaning) representations to inform the correct response. An improved LDT design uses exception word targets (e.g., *yacht*) and pseudohomophone foils which are matched on the semantic and phonological features of the word (e.g., *yawt* is a pseudohomophone for the real word *yacht*). This specialised LDT forces one to rely on orthographic lexical representations because the sublexical assembly of phonology (and its resulting activation of lexical phonology) can no longer be used to successfully determine whether the stimulus spells a real word or not (Ekstrand et al., 2016; Neudorf, Ekstrand, Kress, & Borowsky, 2019; Neudorf, Ekstrand, Kress, Neufeldt, & Borowsky, 2019). For this unique quality of isolating orthographic lexical route processing, we call it an orthographic lexical decision task (OLDT). However, word and bigram frequency effects have not been studied in the context of the specialised OLDT, leaving a gap in the literature regarding whether this lexical decision variant can reflect orthographic lexical processing more clearly than tasks using the nonspecific word stimuli and unmatched nonword foils.

The current study

The goal of the current study is to investigate whether A-O exercise can have an immediate impact on visual processing and reading, such as during a single experiment session. One goal of our first experiment's design was to test whether visual processing tasks in general, or reading tasks specifically, can benefit from A-O exercise. Inspired by the use of video games to improve reading ability, we developed a new type of visual stimulus and task design for use in A-O exercise—a novel health bar decision task (HBDT) which uses stimuli resembling the health bars found in video games in a decision task similar to the OLDT. In the HBDT, the participant indicates whether the health bar matches the number presented. This task will not utilise reading processes as heavily as the OLDT, so it reflects general visual processing. Because it uses stimuli resembling those found in video games, it may better reflect the benefits of video game play on reading in a more controlled environment than what is found in commercial games. The location of the stimuli presented in our two tasks can be manipulated to induce A-O exercise through peripheral or central visual demands.

If standard task practice effects are present (which we refer to as the task-learning hypothesis), participants will experience performance benefits between a first block and a second block, regardless of whether peripheral or central visual demands are presented first. If exposure to peripheral visual demands is beneficial and exposure to central visual demands is detrimental, as our previous reading aloud research suggests (Kress et al., 2023), then participants who are exposed to peripheral before central conditions should show improved performance compared with participants who are exposed to central before peripheral conditions (which we refer to as the A-O exercise hypothesis).

Thus, there are three hypothesised patterns of results we may expect to see in our single-session paradigm:

1. The task-learning hypothesis: If participants' performance is based on general experience with the task, we will expect to see typical task-learning effects, whereby performance in the second block is better than in the first block (see Figure 1a).
2. The A-O exercise hypothesis: If our paradigm promotes A-O exercise in the same manner as previous studies on visual demands in video games, we expect to see a similar dissociation in performance between central and peripheral blocks, as was observed by Kress et al. (2023). In this case, there should be a beneficial effect when peripheral blocks are presented first and a detrimental effect when the central blocks are presented first (see Figure 1b).
3. A blend of task learning and A-O exercise: Both patterns could be present and result in a blended response pattern in our data (see Figure 1c).

In addition, in our OLDT, we will test correlations of performance with the lexical measure of word frequency and the sublexical measure of bigram frequency to confirm that targets in our OLDT are successfully isolating lexical processes, as we expect. In tasks that isolate lexical processing, researchers typically observe a word frequency effect in which performance is better for items with higher word frequency (Borowsky et al., 2013). We expect to

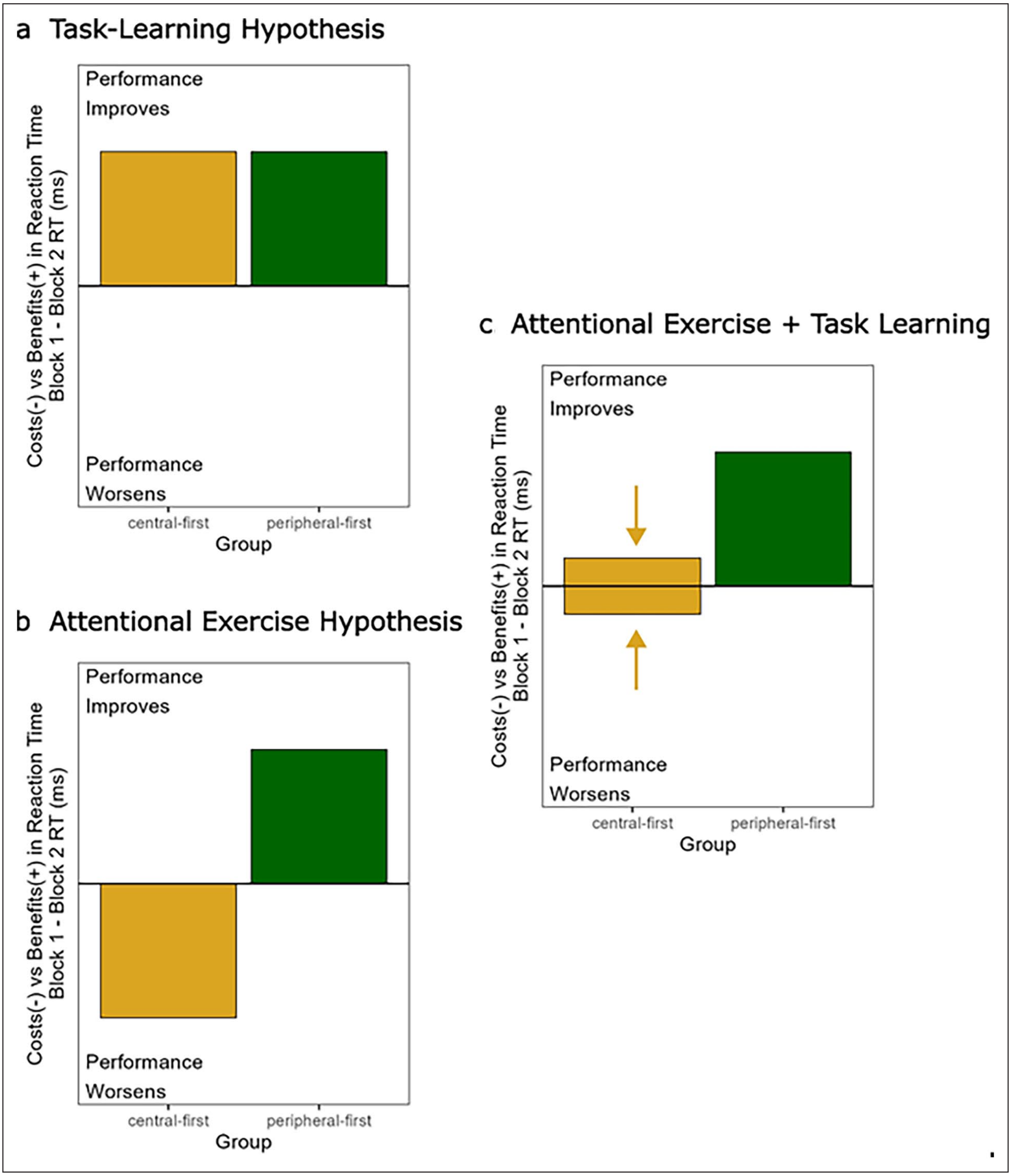

Figure 1. Task learning (a), A-O exercise (b), and A-O exercise + task learning (c) hypotheses with difference scores.
Note. The arrows in the A-O exercise + task-learning hypothesis (c) represent how the combination of task-learning and A-O exercise may lead to minimal change in the central-first group.

replicate this typically observed word frequency effect in our study. Bigram frequency is less well studied, but it has been associated with sublexical processes (Borowsky et al., 2013). Our inclusion of word and bigram frequency analyses will help researchers better understand the role of the lexical and sublexical processes in the OLDT.

Experiment 1

Methods

Participants. This study was approved by the University of Saskatchewan Behavioural Research Ethics Board. We recruited 61 participants[1] (28 cisgender men, 29 cisgender

Table 1. Demographic characteristics as a function of group assignment.

		Central-first group	Peripheral-first group	
Mean age in years (*SD*)		37.63 (13.58)	38.34 (12.34)	$p = .830$
Gender	Cisgender man	15	13	$\chi^2 = 0.14$
	Cisgender woman	16	13	$\chi^2 = 0.31$
	Nonbinary/transgender	1	3	$\chi^2 = 1$

women, 4 nonbinary or transgender participants; $M = 37.97$ years, $SD = 12.90$ years, *mean video game experience* = 8.86 hr/week, $SD = 9.42$, see Table 1) through the Prolific recruitment platform (https://prolific.co) who were covertly prescreened using Prolific's prescreening filter options. Participants were prescreened for country of residence (Canada or the United States), first language (English), language-related disorders (none), vision (normal or corrected-to-normal), and handedness (right). Participants were recruited between November and December 2022. As compensation for participating, participants received £6.75 (Prolific is based in the United Kingdom and provides participants compensation in British Pounds). Participants provided informed consent by registering for the study on Prolific after reading the description (which contained the full written consent form). Informed consent was confirmed by repeating a brief version of the consent form in the PsychoPy experiment, which participants had to agree to with a keypress before the experiment would proceed.

Apparatus and stimuli. The experiment was designed in PsychoPy Version 2022.2.4 (Peirce et al., 2019; *PsychoPy*, 2022) and hosted online through Pavlovia (https://pavlovia.org/). Participants were requested to use a computer with a screen size of 14 inches or larger so that the stimuli and instructions would appear on-screen correctly. In addition, participants used standardised credit card dimensions to calibrate their screen size at the start of the experiment (code adapted from Morys-Carter, 2020) and were instructed to sit approximately 70 cm from their computer monitor to maximise the consistency in the size and visual angle of stimuli between participants.

Peripheral and central presentation. The two tasks in this study were both presented in two different versions that differed in target location. In both versions, a star-shaped visual-spatial cue appeared on-screen for 150 ms in the target location before the stimulus appeared. In the centrally presented conditions, the target location was always the centre of the screen. In the peripherally presented conditions, the target location was randomly selected from one of four locations (top, bottom, left, or right) which were 6.1 cm (4.98 degrees of visual angle using a sitting distance of 70 cm) from the centre of the screen. The stimulus appeared at the target location for 1,000 ms and participants could respond from the moment the stimulus appeared to 500 ms after the stimulus disappeared. If the participant responded before the time was up, then the next trial would begin. If the participant did not respond, then the experiment automatically proceeded to the next trial. The task versions were fully within-subjects (i.e., all participants experienced all conditions), but the order of task presentation was between-subjects.

OLDT. The OLDT is a reading task that focuses on orthographic lexical processes through its use of exception word targets and pseudohomophone foils. The OLDT blocks consisted of 120 exception words (target stimuli; e.g., *ninth*; mean $\log_e$-transformed HAL word frequency = 9.67, $SD = 2.30$; mean bigram frequency = 3,014.44, $SD = 1{,}334.81$) which require orthographic lexical processing to be decoded and 120 corresponding pseudohomophones (foil stimuli; e.g., *nynth*; mean bigram frequency = 2,296.45, $SD = 1{,}650.25$) which do not use orthographic lexical processing but still can activate phonological and semantic processing via phonetic decoding processes. We selected the exception word stimuli from Patterson and Hodges' list of 126 monosyllabic exception words (Patterson & Hodges, 1992) and generated pseudohomophones by changing a letter or multiple letters within the word to disrupt the spelling as much as possible but maintain the same pronunciation (based on Canadian English pronunciations of the words). These exception word and pseudohomophone stimuli are listed in the online Supplementary Material A (Tables A.1 and A.2). The OLDT stimuli were presented in white text (Open Sans font) on a black background, with a maximum word height of 1.1 cm and a maximum word length of 3.5 cm, and flanked by two pairs of white letter x's (i.e., XX),[2] one above and one below the letter-string (making the overall dimensions of these stimuli, including the flankers, 3.8 cm tall by 3.5 cm wide).

HBDT. The HBDT is a novel task designed to reflect aspects of video game play by using stimuli found in video games. The HBDT blocks of the experiment consisted of a set of images depicting a red and white bar (meant to resemble the health bar in a typical video game) which was 1 cm tall by 3.4 cm long and flanked with a two-digit number above and below the bar (overall dimensions of the stimuli, including the flankers, were 3.8 cm tall by 3.4 cm

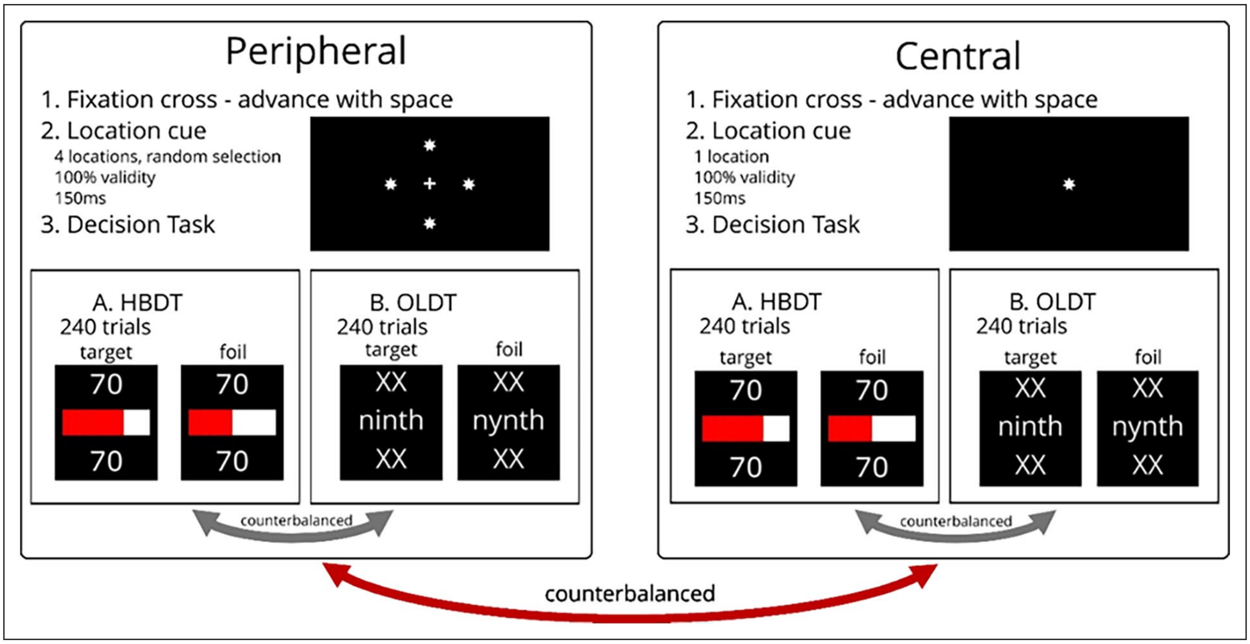

Figure 2. Block counterbalancing and trial procedure for Experiment 1.
Note. The health bar and orthographic LDTs are counterbalanced such that the same task is presented first in both the peripheral and central blocks for a given participant.

wide). The proportion of the bar that was red ranged from 10% to 90%, in increments of 10%. The flanker number ranged from 20 less than the red proportion of the bar to 20 greater than the red proportion of the bar, in increments of 10. The flanking number was restricted such that it could not be lower than 10 or greater than 90, so all flanking numbers were two digits. With these parameters, the set of health bar images totalled 39 unique images—five images each for the health bar values 30 to 70, four images each for the health bar values 20 and 80, and three images each for the health bar values 10 and 90 (see the online Supplementary Material A, Table A.3).

Procedure. After being recruited through Prolific, participants were directed to Pavlovia to complete the online study. The order of the task location blocks was counterbalanced such that half the participants received the centrally presented tasks first. The order of the HBDT and OLDT for each participant was the same in both location (central vs. peripheral) blocks and counterbalanced such that half the participants received the HBDT first:

1. Central-HBDT, 2. Central-OLDT, 3. Peripheral-HBDT, 4. Peripheral-OLDT.
1. Central-OLDT, 2. Central-HBDT, 3. Peripheral-OLDT, 4. Peripheral-HBDT.
1. Peripheral-HBDT, 2. Peripheral-OLDT, 3. Central-HBDT, 4. Central-OLDT.
1. Peripheral-OLDT, 2. Peripheral-HBDT, 3. Central-OLDT, 4. Central-HBDT.

In the HBDT, participants were instructed to press the left arrow key on the keyboard if the proportion of the health bar that was filled red and the number flanking the health bar matched and press the right arrow key if these did not match (240 trials). In the OLDT, participants were instructed to press the left arrow key if the stimulus spelled a real word and press the right arrow key if the stimulus did not spell a real word (240 trials). Figure 2 outlines the counterbalancing and trial process for both tasks. The two tasks would then repeat for the other location block (240 trials × 4 blocks = 960 trials total in the experiment). There was a set of eight practice trials before each block. As part of the instructions, participants were asked to use their index finger to press the left arrow key and their middle finger to press the right arrow key (see Figure 3). The instructed hand position on the keyboard was intended to match the hand position of a similar study where participants responded with mouse clicks (this study is not reported here). After completing the study, participants were redirected to SurveyMonkey (https://www.surveymonkey.com) to complete the postexperiment demographics questions.

Results

For our analyses, we examined by-subjects and by-items median reaction time (which is more robust against outlier responses than mean reaction time) of correct responses to exception word targets and pseudohomophone foils. The values reported in text are the mean of these median

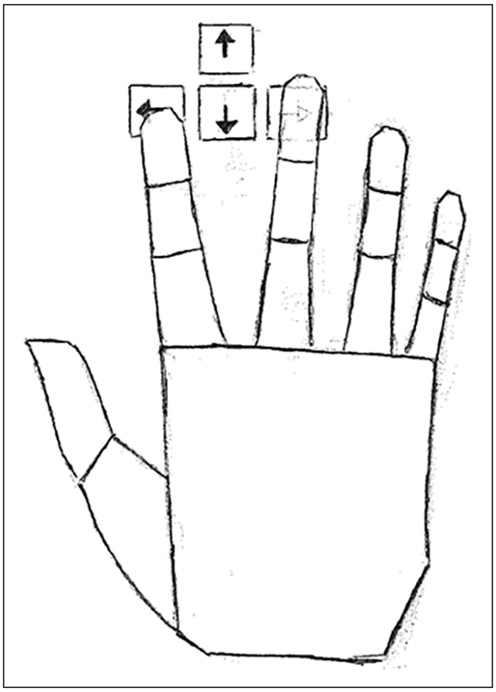

Figure 3. Instructed hand position on arrow keys for Experiment 1.
Note. Participants received written instructions on this hand position.

response times. We also examined by-subjects and by-items mean error rates. Trials with missing responses were removed prior to analysis. OLDT item pairs with outlier error rates (error rates which exceeded 3 *SD* outside the mean error rate; 4 OLDT item pairs (*climb/clime*, *dost/dawst*, *sieve/siv*, and *soot/suht*) were also removed prior to the analysis. For Experiment 1, rather than reporting the by-subjects and by-items analyses together as F_1 and F_2 analyses, respectively, the analyses are reported separately, as the difference scores had to be calculated with distinct formulas for each analysis to handle the counterbalanced design of this study.

By-subjects results. Reaction time and error rate Block 1 − Block 2 difference scores were computed based on the correct responses to target or foil stimuli (see Tables 2 and 3). With this calculation, positive values indicate the performance benefits, whereas negative values indicate the performance detriments.

A set of General Linear Models (GLMs) was used to assess the difference scores in 2 (Task: HBDT vs. OLDT, within-subjects) × 2 (Order: central-first vs. peripheral-first, between-subjects) mixed-measures designs. For exception word targets, the main effect of Task (HBDT: $M=20$ ms, $SD=98$; OLDT: $M=24$ ms, $SD=66$), $F(1, 59)=0.20$, $MSE=4{,}213.69$, $p=.660$, and Task × Order interaction, $F(1, 59)=3.01$, $MSE=4{,}213.69$, $p=.088$, were not significant. There was a main effect of order on target stimuli, $F(1, 59)=36.25$, $MSE=6{,}085.05$, $p<.001$, whereby peripheral-first participants ($M=66$ ms, $SD=56$) had greater improvements in Block 2 than central-first participants ($M=-19$ ms, $SD=83$). We also observed a significant double dissociation of visual demands on OLDT target performance in terms of the 95% confidence intervals, whereby the OLDT target difference scores were greater than zero for peripheral-first participants and less than zero for central-first participants (Masson & Loftus, 2003, see Figure 4).

The same GLM analysis was also conducted on the Block 1 − Block 2 target difference scores for target error rate. There was a significant main effect of task on targets, $F(1, 59)=27.16$, $MSE=68.85$, $p<.001$, whereby the HBDT exhibited larger error rate improvements ($M=7.96\%$, $SD=11.43$) than the OLDT ($M=-0.02\%$, $SD=4.54$). The main effect of Order (central-first: $M=2.71\%$, $SD=10.86$; peripheral-first: $M=5.36\%$, $SD=7.70$) and Task × Order interaction was not significant on target stimuli, $F(1, 59)=2.77$, $MSE=76.87$, $p=.101$ and $F(1, 59)=3.86$, $MSE=68.85$, $p=.054$, respectively.

The same analyses were conducted on the foil stimuli, given our interest in how the stimuli were processed differently in this specialised OLDT. Reaction time and error rate Block 1 − Block 2 difference scores were computed for pseudohomophone foil stimuli (see Table 3) in the same manner as for target stimuli. We used two 2 (Task: HBDT vs. OLDT, within-subjects) × 2 (Order: central-first vs. peripheral-first, between-subjects) mixed-measures GLMs to assess the difference scores.

On reaction time, there was a main effect of Order, $F(1, 59)=33.64$, $MSE=7{,}762.25$, $p<.001$, whereby peripheral-first participants exhibited larger improvements towards foils ($M=73$ ms, $SD=70$) than central-first participants ($M=-20$ ms, $SD=100$). The main effect of Task (HBDT: $M=14$ ms, $SD=120$; LDT: $M=35$ ms, $SD=70$) and Order × Task interaction was not significant on foil stimuli (see Figure 5), all Fs <2.06, all ps $>.157$ ($MSE=7{,}232.79$). On foil error rate, there was a main effect of Order, $F(1, 59)=5.21$, $MSE=42.84$, $p=.026$, whereby peripheral-first participants exhibited larger improvements towards foils ($M=3.57\%$, $SD=6.25$) than central-first participants ($M=0.86\%$, $SD=7.11$). There was also a main effect of Task, $F(1, 59)=22.73$, $MSE=35.21$, $p<.001$, whereby there were larger improvements in the HBDT ($M=4.71\%$, $SD=7.75$) than the OLDT ($M=-0.42\%$, $SD=4.51$). The Order × Task interaction was not significant, $F(1, 59)=.01$, $MSE=35.21$, $p=.905$.

By-items results. With the current design, the perceived benefit of peripheral-first presentation may be due to the peripheral tasks being more difficult than the central tasks, so participants simply perform better in the easier central tasks, regardless of order.

Analysing the data from a by-items perspective allows us to address this issue, as all items were presented in all

Table 2. Experiment 1 reaction time, error rates, and difference scores (by-subjects, on targets).

Order	Task	Block (location)	Mean reaction time ms (*SD*)		Mean error rate % (*SD*)	
Central-first	HBDT	1 (Central)	738	(127)	24.65	(15.38)
		2 (Peripheral)	749	(87)	16.54	(8.44)
		Difference score	−11	(112)	8.11	(12.87)
	OLDT	1 (Central)	609	(70)	8.11	(4.49)
		2 (Peripheral)	636	(84)	10.80	(4.84)
		Difference score	−27	(38)	−2.69	(3.74)
Peripheral-first	HBDT	1 (Peripheral)	761	(99)	23.48	(15.82)
		2 (Central)	707	(97)	15.68	(10.29)
		Difference score	54	(67)	7.80	(9.83)
	OLDT	1 (Peripheral)	653	(102)	8.92	(3.45)
		2 (Central)	574	(76)	6.00	(3.56)
		Difference score	79	(41)	2.92	(3.42)

Table 3. Experiment 1 reaction time, error rates, and difference scores (by-subjects, on foils).

Order	Task	Block (location)	Mean reaction time ms (*SD*)		Mean error rate % (*SD*)	
Central-first	HBDT	1 (Central)	766	(138)	29.90	(12.48)
		2 (Peripheral)	788	(105)	26.55	(13.40)
		Difference score	−22	(49)	0.94	(3.68)
	OLDT	1 (Central)	647	(74)	10.30	(7.45)
		2 (Peripheral)	665	(91)	11.94	(8.38)
		Difference score	−18	(34)	−1.64	(4.89)
Peripheral-first	HBDT	1 (Peripheral)	782	(133)	27.60	(14.85)
		2 (Central)	730	(121)	21.40	(15.44)
		Difference score	52	(82)	6.19	(7.18)
	OLDT	1 (Peripheral)	689	(125)	7.87	(5.15)
		2 (Central)	596	(87)	6.93	(4.17)
		Difference score	93	(49)	0.94	(3.68)

conditions. With this perspective, a Peripheral Attentional Exercise Score is calculated as peripheral-first participants' central block (Central Block 2) minus central-first participants' central block (Central Block 1). The difference score between these two blocks can be considered a view of the peripheral attentional exercise effect, with any effects of difficulty accounted for (because we are comparing the same location). Similarly, a Central Attentional Exercise Score is calculated as central-first participants' peripheral block (Peripheral Block 2) minus peripheral-first participants' peripheral block (Peripheral Block 1).

Reaction time Block 1 − Block 2 difference scores were computed based on the correct responses to target or foil stimuli (see Tables 4 and 5, for reaction times and error rates).

A one-way GLM analysis was conducted on the Block 1 − Block 2 difference scores for target item reaction time. There was a significant main effect of Exercise Type, $F(1, 115) = 12.98$, $MSE = 1{,}462.69$, $p < .001$, whereby peripheral exercise was associated with greater target reaction time improvements ($M = 35$ ms, $SD = 38$) than central exercise ($M = 17$ ms, $SD = 39$). Exercise type was also significant for error rate on targets, $F(1, 115) = 23.18$, $MSE = 37.48$, $p < .001$, with peripheral exercise exhibiting larger target error rate improvements ($M = 2.13\%$, $SD = 5.80$) than central exercise ($M = -1.74\%$, $SD = 6.40$).

As in the previous analyses, reaction time and error rate Block 1 − Block 2 difference scores were computed based on the correct responses to foil stimuli (see Table 5). We used two one-way (Exercise Type: central vs. peripheral) repeated-measures GLMs to assess the difference scores. On foil reaction time, there was a main effect of exercise type, $F(1, 114) = 56.21$, $MSE = 1{,}646.49$, $p < .001$, whereby peripheral exercise exhibits greater improvements on foil reaction time ($M = 53$ ms, $SD = 35$) than central exercise ($M = 13$ ms, $SD = 44$). On foil error rate, there was also a main effect of exercise type, $F(1, 114) = 65.94$, $MSE = 52.30$, $p < .001$, whereby peripheral exercise exhibits greater improvements on foil error rate ($M = 3.34\%$, $SD = 5.27$) than central exercise ($M = -4.12\%$, $SD = 7.96$).

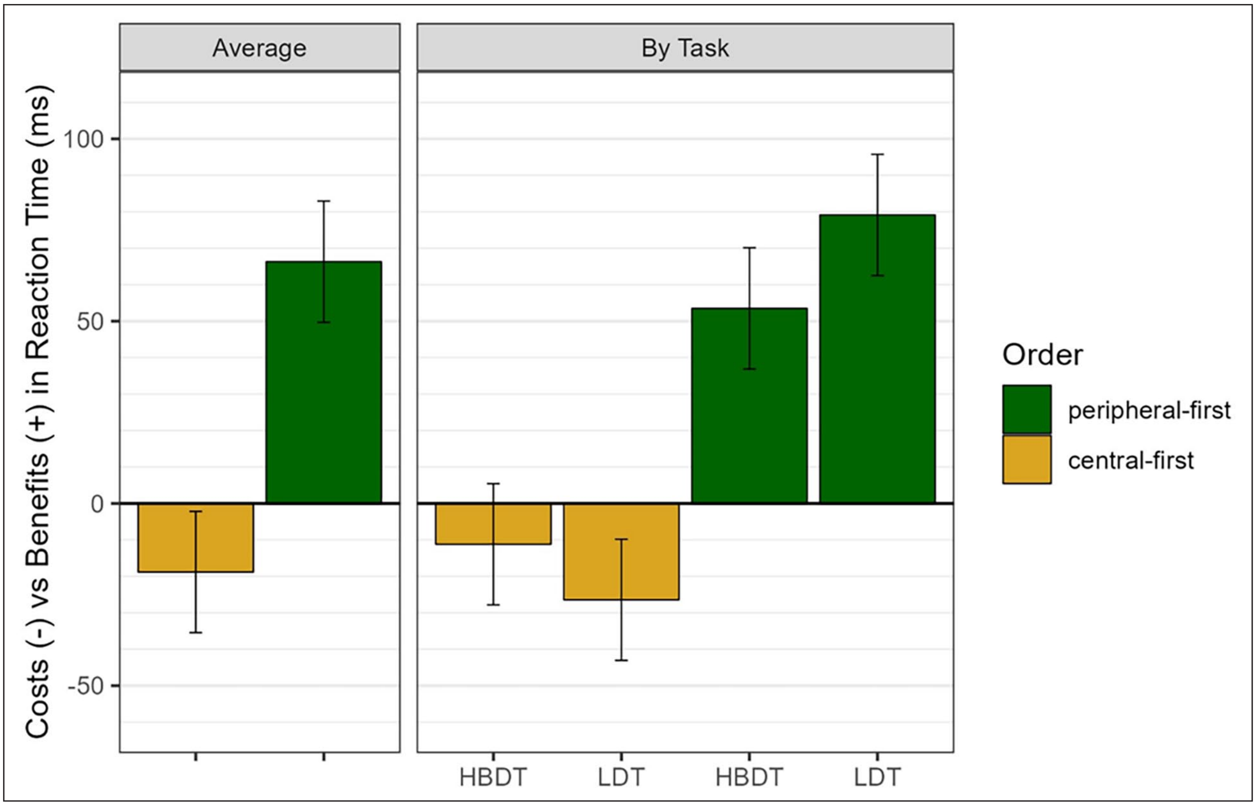

Figure 4. Experiment 1 costs vs. benefits in reaction time as a function of Task and Order (by-subjects, on targets).
Note. Error bars are 95% confidence intervals based on the calculation methods recommended by Masson and Loftus (2003). Positive values indicate performance improved in Block 2.

Word frequency and bigram frequency. To test whether our OLDT is successfully isolating lexical processes, we used simple correlations to analyse the relationship between word frequency, bigram frequency, and our two dependent variables (reaction time and error rate). We used $\log_e$-transformed HAL word frequency as our word frequency measure and mean bigram frequency as our bigram frequency measure (see the online Supplementary Material A). In a previous reading aloud study (Borowsky et al., 2013), these word and bigram frequency measures were used to reflect ventral-lexical and dorsal-sublexical processing, respectively. Log_e-transformed HAL word frequency and mean bigram frequency were derived from the complete English Lexicon Project database (Balota et al., 2007). Pseudohomophones, being nonwords, do not have a word frequency measure associated with them, and as such, we used the base word frequency for these stimuli instead (i.e., the word frequency of the real word from which the pseudohomophone is derived). In our set of stimuli, exception word frequency was not correlated with exception word bigram frequency, $t(118)=-0.04$, $r=-.003$, $p=.97$, nor with pseudohomophone bigram frequency, $t(118)=-0.92$, $r=-.08$, $p=.362$, suggesting that our word and bigram frequency measures are independent.

We observed a significant correlation between word frequency and exception word performance, such that higher word frequencies were associated with faster reaction times and lower error rates (see Table 6). In exception words, there was no significant correlation with bigram frequency and performance. In contrast, we observed a significant correlation between bigram frequency and pseudohomophone performance such that higher bigram frequencies were associated with slower reaction times and higher error rates (see Table 6). There was no significant correlation with base word frequency and performance for the pseudohomophones.

Interim discussion

The results of Experiment 1 show that short-term A-O exercise in the form of exposure to peripheral visual demands may benefit visual processing tasks, and the consistent trends in the word target reaction times and error rates (both in terms of block values and difference scores) suggest there are no speed-accuracy trade-offs. The pattern of results appears to support the blended attentional exercise + task-learning hypothesis, as we see improvements after peripheral blocks, whereas there is minimal change after central blocks (controlling for task learning via by-items analysis). Clearly, as demonstrated by the by-items analysis, a general learning effect is insufficient to explain the results of this experiment.

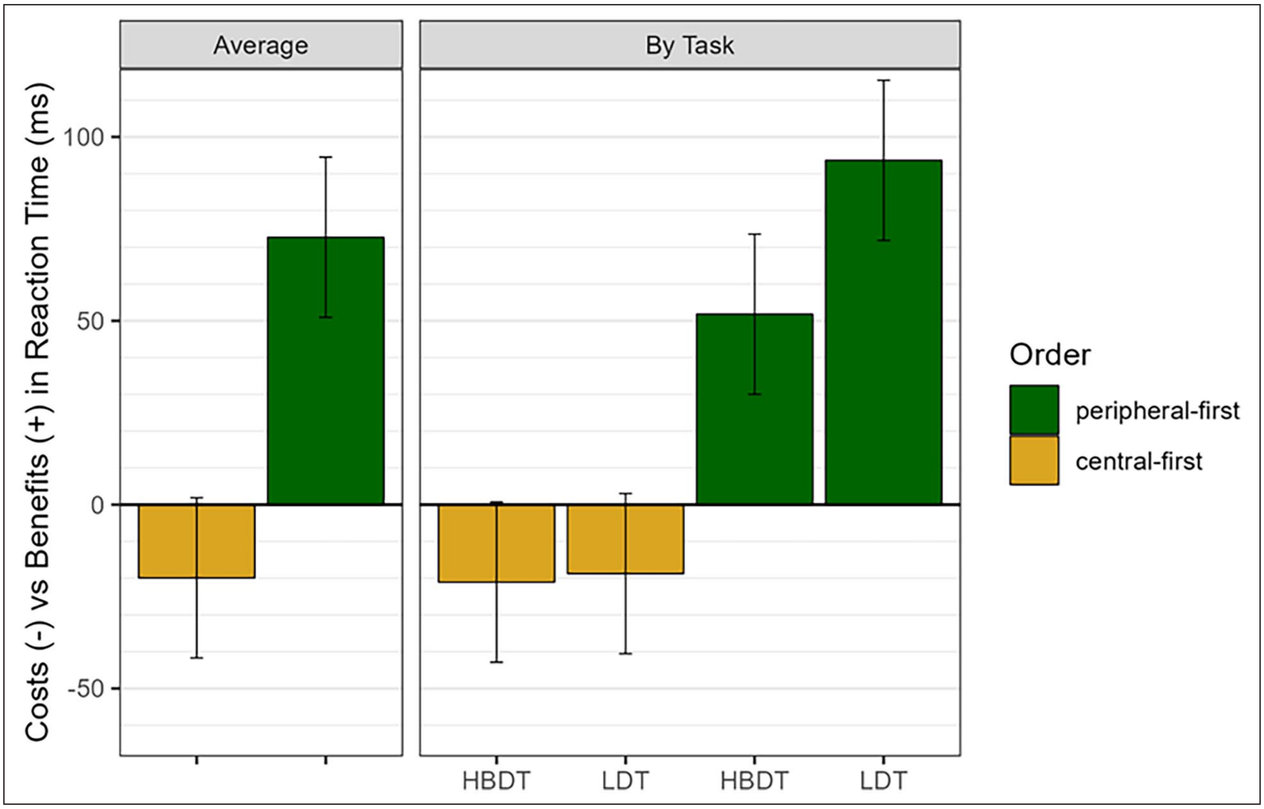

Figure 5. Experiment 1 costs vs. benefits in reaction time as a function of Task and Order (by-subjects, on foils).
Note. Error bars are 95% confidence intervals based on the calculation methods recommended by Masson and Loftus (2003). Positive values indicate the performance improved in Block 2.

Table 4. Experiment 1 reaction time, error rates, and difference scores (by-items, on targets).

Exercise type	Task	Block (participant group's location)	Mean reaction time ms (*SD*)		Mean error rate % (*SD*)	
Central	OLDT	1 (Peripheral-first's peripheral—pretraining)	655	(45)	8.93	(9.96)
		2 (Central-first's peripheral—after central training)	637	(38)	10.67	(10.58)
		Difference score	18	(39)	−1.74	(6.40)
Peripheral	OLDT	1 (Central-first's central—pretraining)	614	(49)	8.13	(10.52)
		2 (Peripheral-first's central—after peripheral training)	579	(33)	6.00	(9.11)
		Difference score	35	(38)	2.13	(5.80)

The results of Experiment 1 demonstrate a clear double dissociation between the higher word frequency and improved performance for exception words and higher bigram frequency and worse performance for pseudohomophones. For pseudohomophones to exhibit worse performance as a function of increased bigram frequency, it suggests that—due to increased familiarity and efficiency in sublexical grapheme-to-phoneme processing—those high-frequency bigrams may be making the pseudohomophones seem more word-like. This evidence supports our argument that our OLDT with exception word targets and pseudohomophone foils is well designed to isolate orthographic lexical processing, where strong word frequency effects support the idea that exception word targets optimally activate orthographic lexical route representations, and bigram frequency may be a useful measure to assess the involvement of the sublexical route for pseudohomophone foil processing in this task. This novel finding is an exciting contribution to the field of lexical decision but should be replicated in additional OLDT studies.

It is important to note that the counterbalance of our design was intended to examine whether reading

Table 5. Experiment 1 reaction time, error rates, and difference scores (by-items, on foils).

Exercise type	Task	Block (participant group's location)	Mean reaction time ms (*SD*)		Mean error rate % (*SD*)	
Central	OLDT	1 (Peripheral-first's peripheral—pretraining)	682	(54)	7.87	(9.41)
		2 (Central-first's peripheral—after central training)	669	(43)	11.99	(11.32)
		Difference score	17	(39)	−1.74	(6.40)
Peripheral	OLDT	1 (Central-first's central—pretraining)	652	(50)	10.27	(11.57)
		2 (Peripheral-first's central block—after peripheral training)	599	(45)	6.93	(9.21)
		Difference score	35	(38)	2.13	(5.80)

Table 6. Simple correlations between reaction time or error rate and word/bigram frequency for all three experiments (targets and foils).

EWs: yacht		Experiment 1 (*df* = 115)	Experiment 2 (*df* = 114)	Experiment 3 (*df* = 113)
Reaction time	WF	−0.64***	−0.64***	−0.65***
	BF	0.02	0.04	0.06
Error rate	WF	−0.54***	−0.54***	−0.52***
	BF	0.02	0.07	0.02
PHs: yawt				
Reaction time	WF	−0.04	−0.06	−0.03
	BF	0.37***	0.33***	0.42***
Error rate	WF	−0.12	−0.12	−0.13
	BF	0.40***	0.40***	0.45***

Note. WF: $\log_e$-transformed HAL word frequency; BF: mean bigram frequency.
Significance levels for the aggregate correlations: *<.05; **<.01; ***<.001.

performance (via the OLDT), general visual processing performance (via the HBDT), both, or neither could be affected by manipulations of A-O exercise. Given our interest in reading processes, our follow-up refines the paradigm to target reading processes specifically. These refinements can also modify the manipulation of A-O exercise in specific test blocks to control the attentional difficulty of the test task. To address concerns regarding increased task difficulty in the peripheral block contributing to the observed results, we developed a refined A-O exercise paradigm that controls the location of the pre- and post-exercise tasks. In addition, we used a similar OLDT to the current study, so that we can test the replicability of the word and bigram frequency double dissociation on performance.

Experiment 2

In Experiment 1, we were interested in the presence of any effect of attentional exercise, as this would support our theory that short-term attentional exercise may be able to benefit reading performance. For Experiments 2 and 3, we refined our A-O exercise paradigm to focus exclusively on the OLDT as our target task of interest, with the HBDT as a training task. These new designs will address Experiment 1's issues regarding task difficulty, as the pre- and post-training tasks are matched on location. In addition, our use of the same OLDT from Experiment 1 gives us the opportunity to replicate the double dissociation between the word frequency and bigram frequency effects on exception word target and pseudohomophone foil performance, respectively.

Hypotheses

In this design, we expected to see an interaction between Training Group and Time, where participants who take part in peripheral A-O exercise should exhibit larger performance improvements than the participants who take part in central A-O exercise. In addition, we expected to see a replication of the double dissociation between the word and bigram frequency effects on exception word and pseudohomophone performance, respectively, which would further support the OLDT as a task to isolate orthographic processing.

Table 7. Demographic characteristics as a function of group assignment.

		Central group	Peripheral group	
Mean age in years (*SD*)		20.65 (4.40)	19.05 (2.26)	$p=.159$
Gender	Cisgender man	4	5	$\chi^2=0.11$
	Cisgender woman	16	15	$\chi^2=0.03$

Note. All participants reported using the mouse with their right hand.

Method

Participants. We recruited 40 participants (9 cisgender men, 31 cisgender women; $M=19.85$ years, $SD=3.54$, *mean video game experience*=2.40 hr/week, $SD=4.40$) through the University of Saskatchewan SONA participant pool during the Fall 2023 term. As compensation for participating, participants received two bonus course credits. The first language of all participants was English. In addition, although two participants were left-handed, all participants reported using the mouse with their right hand. This study was approved by the University of Saskatchewan Research Ethics Board, and all participants provided informed consent before participating.

Participants were randomly assigned to the training groups such that half the participants received the central HBDT as their training task and the other half received the peripheral HBDT as their training task. There were no significant differences between the training groups in terms of age or gender (see Table 7).

Apparatus and stimuli. The experiment was designed in PsychoPy Version 2023.1.3 (Peirce et al., 2019; *PsychoPy*, 2023) and run in a dimly lit, quiet experiment room. The experiment computer was a Lenovo ThinkCentre M70q computer running Windows 10. An HP E243 23.8-inch monitor displayed the experiment stimuli. Responses were recorded with a standard keyboard and mouse. The OLDT and HBDT stimuli were the same as in Experiment 1. The OLDT stimuli were split into two half-blocks (see the online Supplementary Material A). These half-blocks were counterbalanced such that half the participants in each training group received Half-Block A as the pretraining task and the other participants received Half-Block B as the pretraining task.

Procedure. An experimenter was present for the duration of the experiment, and there was a set of eight practice trials before each of the three tasks. After providing informed consent, participants began the experiment with the pretraining reading task, which was a 120-trial OLDT (half the length of the OLDT in Experiment 1). In the half-OLDT, a fixation cross would appear, and participants would press the spacebar to begin the trial. After a 150-ms interstimulus interval, where no fixation cross, stimulus, or cue was on-screen, a letter-string stimulus would appear. Participants were instructed to left-click if the letter-string spelled a real word, and right-click if the letter-string did not spell a real word. The OLDT stimuli were presented until response. Then participants received one block (240 trials) of either the central or peripheral HBDT as their training task, depending on the group to which they were assigned. The HBDT followed the same timing and procedure as in Experiment 1 with one exception. Instead of using the keyboard, participants were instructed to use the mouse to left-click for target responses and right-click for foil responses. After the training task, participants completed a post-training reading task, which was a half-OLDT following the procedure of the pretraining task but using the other set of 120 trials. Finally, participants completed their demographics form and were debriefed on the purpose of the study. Figure 6 depicts the design and trial procedure of this experiment.

Results

As in Experiment 1, we examined participant (reported as F_1) or item (reported as F_2) median reaction time and mean error rate to correct responses on both exception word targets and pseudohomophone foils. The reaction times reported in the text are the mean of these median response times. OLDT item pairs with outlier error rates (error rates which exceeded 3 *SD* outside the mean error rate) were removed prior to the analysis (six item pairs: *brooch/broatch*, *dost/dawst*, *sieve/siv*, *soot/suht*, *suave/swawv*, *suede/swaid)*. Linear mixed model and general linear mixed model analyses were conducted in R (R Core Team 2022), using the methods of Baayen et al. (2008), Bates et al. (2015), and Matuschek et al. (2017), and can be found in the online Supplementary Material B.

Target stimuli analyses. First, 2 (Training Group: central vs. peripheral) × 2 (Time: pretraining vs. post-training) mixed-measures GLMs were conducted on median reaction times to target stimuli to test the effect of A-O exercise (through the HBDT) on OLDT performance. There was a consistent main effect of Time, $F_1(1, 38)=20.92$, $MSE_1=1{,}878.54$, $p_1<.001$, $F_2(1, 113)=75.33$, $MSE_2=3{,}305.54$, $p_2<.001$, whereby responses to targets were faster post-training ($M_1=581$ ms, $SD_1=69$, $M_2=595$ ms, $SD_2=75$) than pretraining ($M_1=625$ ms,

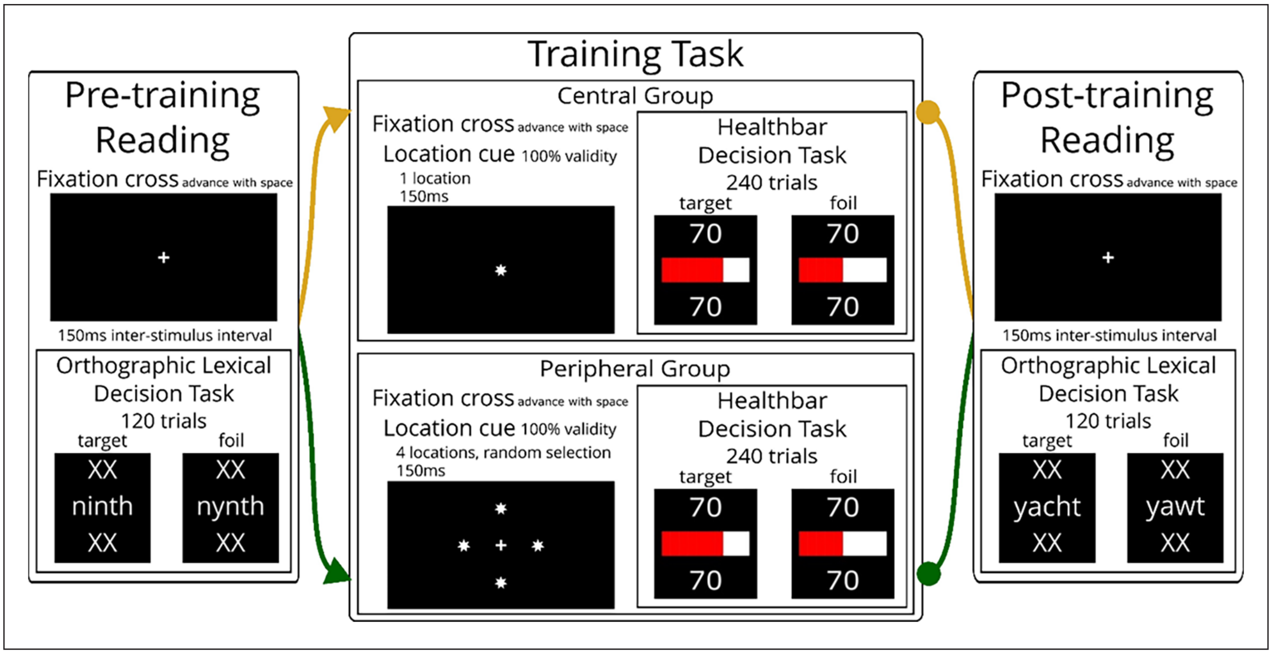

Figure 6. Training group assignment and trial procedure for Experiment 2.
Note. A similar procedure was used for Experiment 3, although the pre- and post-training reading tasks were peripherally presented (un-cued), and the training task was doubled in length (two blocks of 240 trials, with a brief break in-between).

SD_1=85, M_2=642 ms, SD_2=86). There was no effect of Training Group (peripheral: M_1=606 ms, SD_1=87, M_2=620 ms, SD_2=85; central: M_1=601 ms, SD_1=74, M_2=616 ms, SD_2=83), $F_1(1, 38)=0.05$, $MSE_1=10{,}444.53$, $p_1=.820$, $F_2(1, 113)=.70$, $MSE_2=2{,}443.64$, $p_2=.403$. There also was no Training Group × Time interaction, $F_1(1, 38)=0.003$, $MSE_1=1{,}878.54$, $p_1=.958$ (see Figure 7), $F_2(1, 113)=.61$, $MSE_2=2{,}000.89$, $p_2=.436$.

The same GLMs were conducted on mean error rates to targets. There was a consistent main effect of Training Group on targets, $F_1(1, 38)=6.29$, $MSE_1=38.74$, $p_1=.017$, $F_2(1, 113)=24.23$, $MSE_2=57.93$, $p_2<.001$, whereby the peripheral training group had worse error rates (M_1=9.77%, SD_1=6.68; M_2=9.82%, SD_2=14.23) than the central training group (M_1=6.28%, SD_1=3.71; M_2=6.32%, SD_2=12.51). The main effect of Time (pretraining: M_1=7.57%, SD_1=5.34, M_2=7.59%, SD_2=13.04; post-training: M_1=8.48%, SD_1=5.97, M_2=8.55%, SD_2=13.96) was not significant $F_1(1, 38)=0.82$, $MSE_1=20.31$, $p_1=.372$, $F_2(1, 113)=1.33$, $MSE_2=80.03$, $p_2=.252$. The Training Group × Time interaction was also not significant, $F_1(1, 38)=0.82$, $MSE_1=20.31$, $p_1=.372$, $F_2(1, 113)=1.34$, $MSE_2=79.15$, $p_2=.249$.

Foil stimuli analyses. The same 2 (Training Group: central vs. peripheral) × 2 (Time: pretraining vs. post-training) mixed-measures GLMs were conducted on median reaction times to foil stimuli. There was a main effect of Time on foils (see Figure 8), $F_1(1, 38)=38.38$, $MSE_1=3{,}751.48$, $p_1<.001$, $F_2(1, 113)=144.14$, $MSE_2=3{,}681.23$, $p_2<.001$, whereby responses were faster after training (M_1=651 ms, SD_1=76; M_2=665 ms, SD_2=87) than before training (M_1=735 ms, SD_1=119; M_2=733 ms, SD_2=101). There was no effect of Training Group by-subjects (central: M_1=699 ms, SD_1=108; peripheral: M_1=687 ms, SD_1=109), $F_1(1, 38)=0.16$, $MSE_1=16{,}594.04$, $p_1=.696$. However, the effect of Training Group was significant by-items, $F_2(1, 113)=12.02$, $MSE_2=3{,}425.37$, $p_2<.001$, whereby foils were responded to more quickly by participants in the peripherally trained group (M_2=690 ms, SD_2=100) than participants in the centrally trained group (M_2=709 ms, SD_2=99). There was no Training Group × Time interaction, $F_1(1, 38)=1.37$, $MSE_1=3{,}751.48$, $p_1=.250$, $F_2(1, 113)=0.11$, $MSE_2=4{,}294.18$, $p_2=.741$.

On mean error rate for foils, there were no by-subjects effects or interactions, all $F_1\text{s}<1.01$, all $p_1\text{s}>.323$ (Time-pretraining: M_1=13.23%, SD_1=8.93; Time-post-training: M_1=13.57%, SD_1=9.04; Training Group-central: M_1=13.19%, SD_1=10.23; Training Group-peripheral: M_1=13.61%, SD_1=7.55). There were also no by-items effects or interactions, all $F_2\text{s}<2.38$, all $p_2\text{s}>.126$ (Time-pretraining: M_2=13.33%, SD_2=15.23; Time-post-training: M_2=13.60%, SD_2=15.91; Training Group-central: M_2=13.20%, SD_2=15.56; Training Group-peripheral: M_2=13.73%, SD_2=15.58).

Word frequency and bigram frequency. As in Experiment 1, we used correlations to analyse the relationship between word frequency, bigram frequency, and our two dependent variables (reaction time and error rate). We used the same

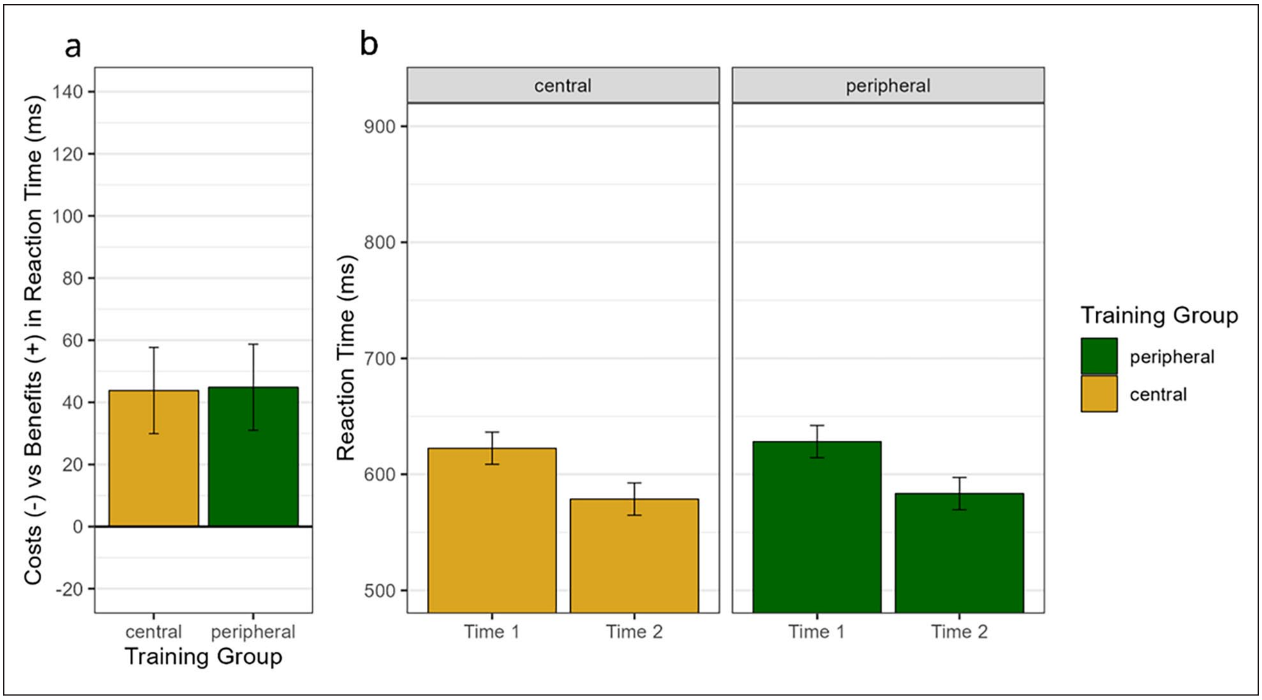

Figure 7. Experiment 2 reaction time as a function of time and training group (by-subjects, on targets) depicted (a) as difference scores and (b) by Time.

Note. Error bars are 95% confidence intervals based on the calculation methods recommended by Masson and Loftus (2003).

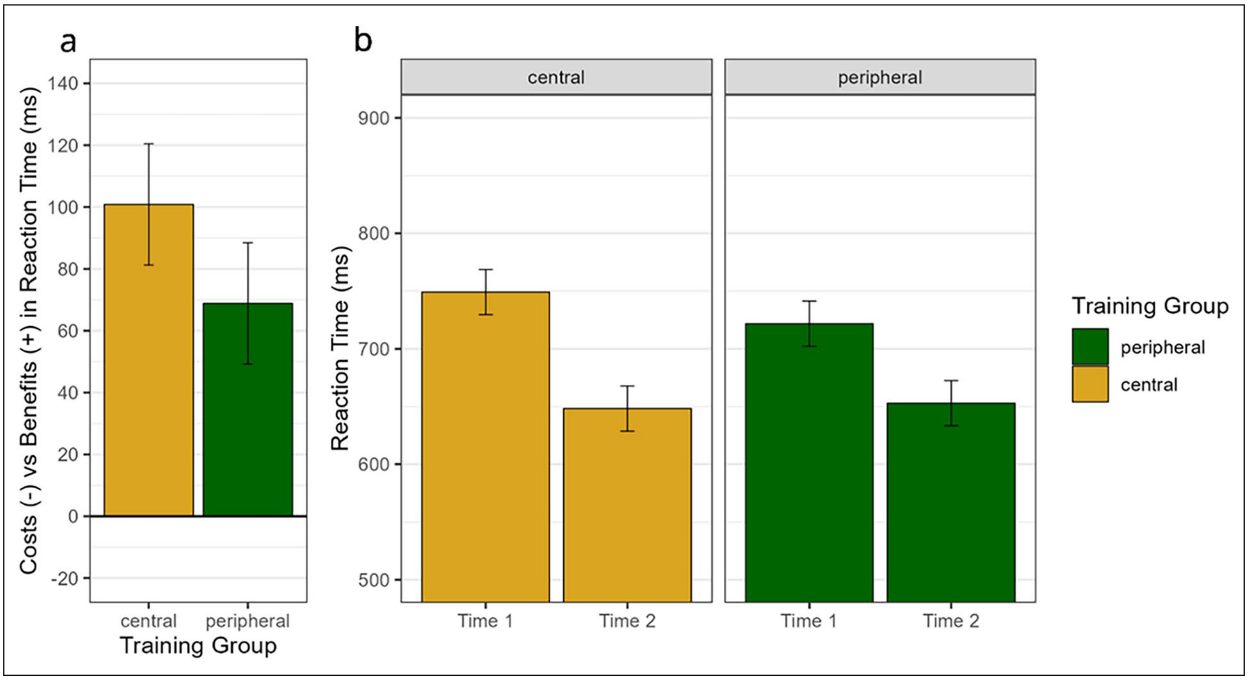

Figure 8. Experiment 2 reaction time as a function of time and training group (by-subjects, on foils) depicted (a) as difference scores and (b) by Time.

Note. Error bars are 95% confidence intervals based on the calculation methods recommended by Masson and Loftus (2003).

word and bigram frequency measures as in Experiment 1. The word and bigram frequency results were consistent with Experiment 1 (see Table 6). Word frequency had a negative correlation with exception word reaction times and error rate (i.e., improved performance as a function of increased word frequency), whereas bigram frequency had

Table 8. Demographic characteristics as a function of group assignment.

		Central group	Peripheral group	
Mean age in years (*SD*)		21.75 (5.79)	23.84 (7.44)	$p = .332$
Gender	Cisgender man	3	9	$\chi^2 = 3.00$
	Cisgender woman	15	10	$\chi^2 = 1.00$
	Nonbinary/transgender	1	0	
	Prefer not to respond	1	0	

Note. All participants reported using the mouse with their right hand.

a positive correlation with pseudohomophone reaction time and error rate (i.e., worse performance as a function of increased bigram frequency).

Interim discussion

With Experiment 2, we were able to replicate the double dissociation of a word frequency effect on exception word performance, a bigram frequency effect on pseudohomophone performance, and the absence of these effects with their counterparts. However, we did not observe a difference between the peripheral and central A-O exercise groups in this refined paradigm. It may be the case that in the new paradigm, the single block of A-O exercise was not enough to induce the additional benefit expected of peripheral A-O exercise. To address this, we extended the exercise duration in Experiment 3 to boost the paradigm's sensitivity for detecting differences in these exercise groups.

Experiment 3

Method

Participants. We recruited 39 participants who were included in the analysis[3] (12 cisgender men, 25 cisgender women, 1 nonbinary or transgender participants, 1 preferred not to respond; $M = 22.77$ years, $SD = 6.64$, *Mean video game experience* = 8.09 hr/week, $SD = 13.18$) through the University of Saskatchewan SONA participant pool during the Winter and Spring 2024 terms. As compensation for participating, participants received two bonus course credits. As in Experiment 2, the first language of all participants was English, and all participants reported using the mouse with their right hand (although four participants were otherwise left-handed). This study was approved by the University of Saskatchewan Research Ethics Board, and all participants provided informed consent before participating.

Participants were randomly assigned to the training groups such that half the participants received the central HBDT as their training task and the other half received the peripheral HBDT as their training task. There were no significant differences between the training groups in terms of age or gender (see Table 8).

Apparatus and stimuli. The apparatus and stimuli in Experiment 3 were the same as those used in Experiment 2.

Changes in procedure. The experiment design and procedure were nearly the same as Experiment 2, with two major exceptions. The first change was in the training HBDT length. The training HBDT was doubled (i.e., two HBDT blocks instead of one), with a brief participant-controlled break of 10–30 s to reduce eye-fatigue during training. The second change was in the OLDT design. The OLDT blocks for Experiment 2 were peripherally presented to increase the difficulty of the task. This also made the design more similar to our earlier experiment where we first observed the double dissociation between central and peripheral visual demands on reading performance (Kress et al., 2023). The peripherally presented OLDT was not cued, to keep the stimulus presentation consistent between Experiments 2 and 3, and was presented until response or for a maximum of 1,500 ms (the HBDT also used this timing). The participant would see the fixation cross and press the spacebar to begin the trial. Then, after a 150-ms interstimulus interval, where no fixation cross, stimulus, or cue was on-screen, the letter-string stimulus would appear randomly in one of the four target locations, and the participant would make their response . The four target locations were the same as those used in the HBDT.

Results

The analysis for Experiment 3 followed the same steps as the previous experiment. We examined participant (F_1) or item (F_2) median reaction time and mean error rate to correct responses and the values reported in text are the mean of these median response times. OLDT item pairs with outlier error rates (error rates which exceeded 3 *SD* outside the mean error rate) were removed prior to the analysis (seven item pairs: *dost/dawst, leapt/lept, mauve/moav, sieve/siv, soot/suht, suave/swawv, suede/swaid*). Linear mixed model and general linear mixed model analyses can be found in the online Supplementary Material C.

Target stimuli analyses. As in the previous experiment, 2 (Training Group: central vs. peripheral) × 2 (Time: pre-training vs. post-training) mixed-measures GLMs were

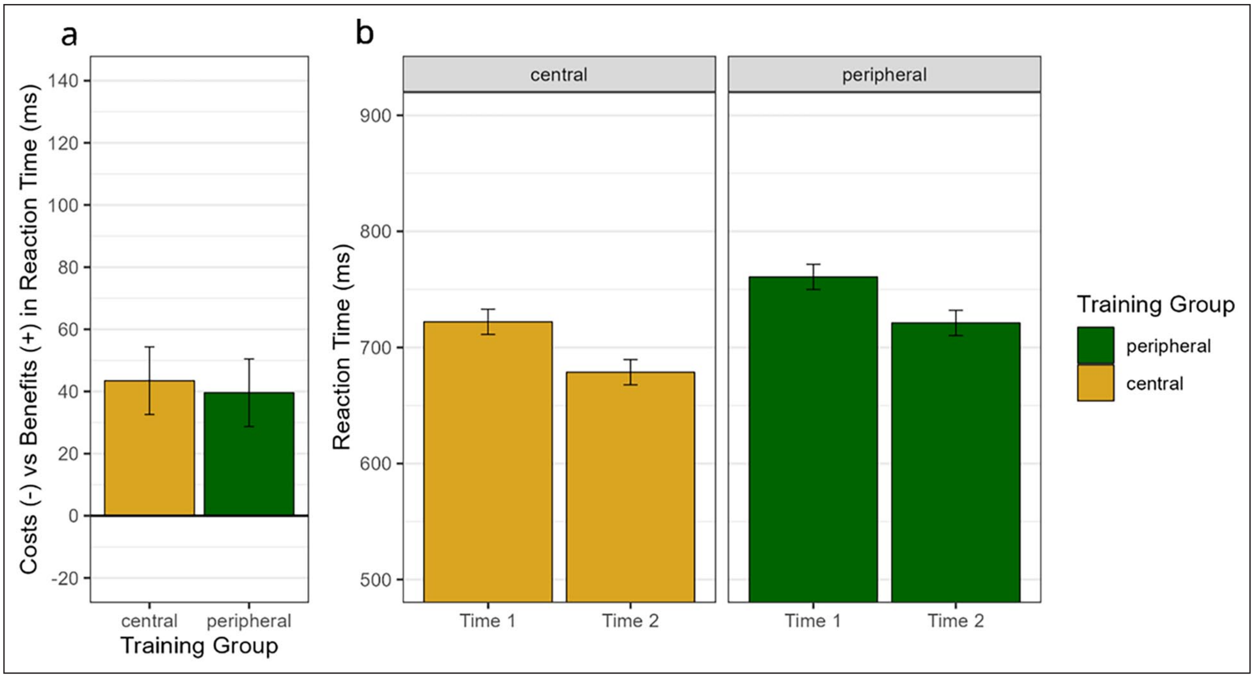

Figure 9. Experiment 3 reaction time as a function of time and training group (by-subjects, on targets) depicted (a) as difference scores and (b) by Time.
Error bars are 95% confidence intervals based on the calculation methods recommended by Masson and Loftus (2003).

conducted on median reaction times to target stimuli to test the effect of attentional exercise (through the HBDT) on OLDT performance. There was a main effect of Time, $F_1(1, 37)=29.91$, $MSE_1=1,123.67$, $p_1<.001$, $F_2(1, 112)=79.74$, $MSE_2=3,574.53$, $p_2<.001$, whereby responses were faster post-training ($M_1=699$ ms, $SD_1=70$; $M_2=711$ ms, $SD_2=72$) than pretraining ($M_1=741$ ms, $SD_1=71$; $M_2=761$ ms, $SD_2=90$). There was also an effect of Training Group (central: $M_1=700$ ms, $SD_1=68$, $M_2=713$ ms, $SD_2=81$; peripheral: $M_1=741$ ms, $SD_1=74$, $M_2=759$ ms, $SD_2=84$), $F_1(1, 37)=3.89$, $MSE_1=8,261.10$, $p_1=.056$, $F_2(1, 112)=82.45$, $MSE_2=2,962.74$, $p_2<.001$. There was no significant Time × Training Group interaction, $F_1(1, 37)=0.06$, $MSE_1=1,123.67$, $p_1=.801$, $F_2(1, 112)=0.45$, $MSE_2=3,605.77$, $p_2=.505$. Figure 9 depicts the by-subjects main effect of Time in both groups.

When the same GLM was conducted on error rates, there was a significant effect of Training Group, $F_1(1, 37)=7.48$, $MSE_1=49.69$, $p_1<.001$, $F_2(1, 112)=37.19$, $MSE_2=59.52$, $p_2<.001$, with participants in the central group ($M_1=6.28\%$, $SD_1=4.69$; $M_2=6.30\%$, $SD_2=10.83$) responding with less errors than participants in the peripheral group ($M_1=10.65\%$, $SD_1=6.50$; $M_2=10.73\%$, $SD_2=14.19$). The main effect of Time was significant by-items but not by-subjects (pretraining: $M_1=7.69\%$, $SD_1=6.02$, $M_2=7.78\%$, $SD_2=11.56$; post-training: $M_1=9.12\%$, $SD_1=6.02$, $M_2=9.25\%$, $SD_2=13.92$), $F_1(1, 37)=2.73$, $MSE_1=14.74$, $p_1=.107$, $F_2(1, 112)=3.94$, $MSE_2=62.33$, $p_2=.050$. The Time × Training Group interaction was not significant, $F_1(1, 37)=0.21$, $MSE_1=14.74$, $p_1=.900$, $F_2(1, 112)=0.006$, $MSE_2=61.13$, $p_2=.938$.

Foil stimuli analyses. We also ran the 2 (Training Group: central vs. peripheral) × 2 (Time: pretraining vs. post-training) mixed-measures GLMs on median reaction times to foil stimuli (see Figure 10). There was a main effect of Time, $F_1(1, 37)=48.22$, $MSE_1=2,101.40$, $p_1<.001$, $F_2(1, 112)=154.02$, $MSE_2=3,411.58$, $p_2<.001$, whereby participants responded more quickly to foils after training ($M_1=766$ ms, $SD_1=86$; $M_2=776$ ms, $SD_2=90$) than before training ($M_1=838$ ms, $SD_1=113$; $M_2=844$ ms, $SD_2=92$). There was no effect of Training Group (central: $M_1=798$ ms, $SD_1=103$, $M_2=809$ ms, $SD_2=95$; peripheral: $M_1=806$ ms, $SD_1=111$, $M_2=812$ ms, $SD_2=99$), $F_1(1, 37)=0.06$, $MSE_1=18,596.49$, $p_1=.807$, and there was no Training Group × Time interaction, $F_1(1, 37)=0.03$, $MSE_1=2,101.40$, $p_1=.873$, $F_2(1, 112)=0.69$, $MSE_2=4,442.25$, $p_2=.407$.

On error rate for foils, there was a main effect of Time, $F_1(1, 37)=5.32$, $MSE_1=22.41$, $p_1=.027$, $F_2(1, 112)=6.09$, $MSE_2=99.95$, $p_2=.015$, whereby participants responded with more errors after training ($M_1=14.50\%$, $SD_1=9.20$, $M_2=14.51\%$, $SD_2=17.26$) than before training ($M_1=12.03$, $SD_1=9.41$, $M_2=12.19\%$, $SD_2=14.24$). This speed-accuracy trade-off as a function of time is explored in the online Supplementary Material D with an analysis of incorrect trial reaction times. The effect of Training Group (central: $M_1=14.56\%$, $SD_1=8.54$, $M_2=12.08\%$, $SD_2=14.29$; peripheral: $M_1=11.90\%$, $SD_1=10.03$, $M_2=14.61\%$, $SD_2=17.20$) was not significant by-subjects $F_1(1, 37)=0.91$, $MSE_1=151.75$, $p_1=.347$, but was significant by-items, $F_2(1, 112)=6.25$, $MSE_2=116.31$, $p_2=.014$. The Training Group ×

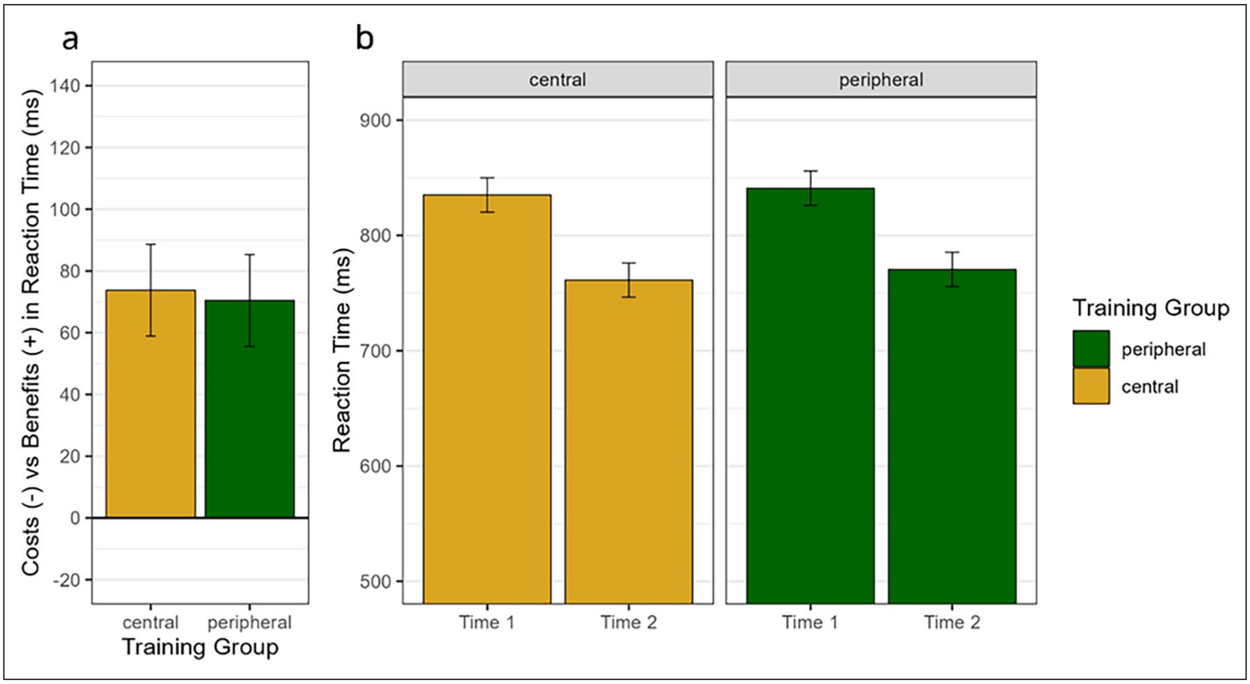

Figure 10. Experiment 3 reaction time as a function of time and training group (by-subjects, on foils) depicted (a) as difference scores and (b) by Time.
Error bars are 95% confidence intervals based on the calculation methods recommended by Masson and Loftus (2003).

Time interaction was not significant, $F_1(1, 37)=0.02$, $MSE_1=22.41$, $p_1=.876$, $F_2(1, 112)=0.004$, $MSE_2=95.19$, $p_2=.945$.

Word frequency and bigram frequency. As in the previous experiments, we used simple correlations to analyse the relationship between word frequency, bigram frequency, and our two dependent variables (reaction time and error rate). We used the same word and bigram frequency measures as the previous experiments and followed the same procedure. Replicating the results of the previous experiments, we observed a significant correlation between word frequency and exception word performance, such that higher word frequencies were associated with faster reaction times and lower error rates, and a significant correlation between bigram frequency and pseudohomophone performance, such that higher bigram frequencies were associated with slower reaction times and worse error rates (see Table 6). As in the previous experiments, bigram frequency was not correlated with exception word performance, and base word frequency was not correlated with pseudohomophone performance.

Interim discussion

Overall, the results of Experiment 3 are consistent with the results of Experiment 2. The same pattern of word and bigram frequency effects was observed, so the double dissociation appears to be robust. There was also an overall main beneficial effect of Time on reaction time, with no interaction, similar to Experiment 2, and suggesting both our central and peripheral training conditions may be help reading performance. However, unlike Experiment 2, this was accompanied by some evidence for a detrimental effect of Time on error rates, suggesting a form of speed-accuracy trade-off or fatigue effect. These findings mean the length of the training period in our third experiment may be too long for a single session, and future studies should consider an intermediate number of trials between the length of Experiments 2 and 3 to maximise the benefits of training before reaching this fatigue effect.

Discussion

Bringing together the research fields of reading and attention, the goal of this set of studies was to test whether short-term attentional exercise could benefit reading performance. Our hypothesis that attentional exercise may benefit reading ability drew from the past literature on cognitive exercise (e.g., Nouchi et al., 2012), past links between reading and attention (Boden & Giaschi, 2007; Ekstrand, Neudorf, Kress, & Borowsky, 2019), and recent work on the relationship between video games and both reading and attention (Bertoni et al., 2021; Kress et al., 2023).

Potential A-O exercise effects

In both the by-subjects and by-items analyses of Experiment 1, we observed an effect of Order, whereby the peripheral A-O exercise demonstrated improved performance over time, whereas central A-O exercise demonstrated either minimal change or worsened performance over time. These reaction time benefits did not come at the cost to accuracy, meaning participants were not experiencing speed-accuracy trade-offs with this improved reaction time performance. This was replicated by-subjects and by-items, demonstrating that the effects could not be attributable to general learning processes.

The designs of Experiments 2 and 3 were intended to refine the paradigm by controlling the location of the pre- and post-training reading task. In this case, we observed a significant effect of Time, whereby all participants' performance improved between pre- and post-training; however, there was no interaction that would suggest peripheral A-O exercise was more effective than central A-O exercise at inducing reading improvements in the OLDT. It may have been the case that Experiment 1's effects were driven by task difficulty, although note that the interleaved design of the blocks in Experiment 1 meant that there was the potential for a longer period for differences to emerge relative to Experiment 2. Furthermore, the significant by-items effect of Task in the Experiment 1 analyses suggests that the effect cannot be fully accounted for by task difficulty alone. Likewise, general learning effects would not account for the findings observed in Experiment 1, so Experiments 2 and 3 must have other factors beyond task learning which drive the improvements observed after both central and peripheral training. The trade-off between speed and accuracy in Experiment 3 also indicates there is more than simple task learning at play in these designs. The next sections will unpack the reading and attentional processes present in our A-O exercise paradigm as we consider explanations for the results of Experiments 2 and 3 and options to further refine the A-O exercise programme.

Which attentional processes benefit from A-O exercise?

Our goal was to test whether short-term A-O exercise had the potential to provide any amount of reading performance benefit, and we can bring past findings together with the current study to make some predictions regarding the relevant systems. Our A-O exercise consisted of briefly flashing a star at the target location, which one could argue would activate both the alerting (when will the target appear) and orienting (where will the target appear) mechanisms of the attentional network (Dye et al., 2009; Fan et al., 2002). In the dyslexia research field, the dysfunction in the orienting system is most frequently discussed when exploring the overlapping reading + attention deficits in dyslexia (Boden & Giaschi, 2007; Gori et al., 2014, 2016; Taran et al., 2022). Oculomotor control and attentional orienting is also more frequently discussed over the alerting system when testing action video games as a cognitive exercise tool (Diarra et al., 2019; Dye et al., 2009; Pasqualotto et al., 2022; West et al., 2013). For these reasons, we suspect that attentional orienting, rather than attentional alerting, is the more relevant process when designing an A-O exercise task to benefit reading. These attentional orienting processes could involve large oculomotor shifts, such as those required to shift from the centre of the screen to the peripheral of the screen during a video game, or from the end of one line to the start of the next line while reading a book. Attentional orienting processes could also involve small oculomotor shifts, such as those required to precisely target the enemy you are looking at in a video game, or those required to shift between, or within, nearby words in a sentence.

When one considers attentional orienting processes, this can be subcategorised into the voluntary and reflexive orienting processes discussed by Corbetta and Shulman (2002). The star cue could be considered a reflexive orienting cue, albeit one with 100% validity. In the peripheral A-O exercise, this reflexive orienting cue required large oculomotor shifts to move from the central fixation cross to the target location for the HBDT. The reflexive orienting cue for the central A-O exercise condition would not require this same large oculomotor shift. It would be reasonable to expect the exercise of this reflexive orienting process to benefit reading, as we originally hypothesised, given that past research has suggested action video games (which emphasise peripheral processing) benefit reading processes (Bertoni et al., 2021; Franceschini et al., 2017; Pasqualotto et al., 2022).

Our designs, which varied the degree of large reflexive orienting movements, did not lead to differences in the reading performance between the central and peripheral A-O exercise groups. This could mean that both versions of our task contained a form of A-O exercise that was beneficial to reading.

Voluntary orienting processes, and in particular fine-grained voluntary oculomotor processes, would be good A-O exercise to consider in future designs. The HBDT involves small voluntary oculomotor movements from the health bar in the centre of the screen, up or down to the flanker number, and then potentially back to the health bar to evaluate its size. These small oculomotor movements are present in both the central and peripheral versions of the task. Reading also involves small systematic oculomotor movements between word segments or whole words, and large oculomotor movements generally only need to take place when you reach the end of the line or the end of the page. With this in mind, it could be the case that both peripheral and central versions of the HBDT were sufficient to induce the type of A-O exercise that could most benefit reading (i.e., small oculomotor movements).

With this new idea of exercising small, rather than large, oculomotor movements, future refinements of the A-O exercise paradigm should include a version of the

HBDT to reduce small oculomotor movements in one of the training groups. This could be done by overlaying the number with the health bar or by briefly presenting the number at the target location, followed by the health bar in the same location. Either of these designs would minimise the involvement of small oculomotor movements in the given group and help determine what type of attentional process is most important for A-O exercise to improve reading. Ideally, a future study would include three conditions, one each for large movements, small movements, and no movements to compare the effectiveness of all three options against the others. In addition, eye-tracking could be employed to verify the magnitude of the oculomotor movements that are occurring in each condition.

Isolating orthographic processing

The reading task we used in this set of experiments was an OLDT, where participants indicated whether the presented letter-string spelled a real word or not. In this case, the foils were pseudohomophones, which are nonwords that sound like a real word if they are phonetically decoded (Cummine et al., 2015; Neudorf, Ekstrand, Kress, Neufeldt, & Borowsky, 2019; Wingerak et al., 2017). The use of exception word targets and pseudohomophone foils should force orthographic lexical processing in tasks such as the OLDT, as an incorrect response will be generated if participants focus on phonological and/or semantic instead of orthographic lexical representations.

Our theory that the OLDT forces orthographic lexical processing is supported by the results in our word and bigram frequency analyses which show that performance is better for exception words that have high word frequency, whereas performance is worse for pseudohomophones that have high bigram frequency. Our word frequency effect aligns with past research that has attributed the effect to enhanced orthographic processing (Borowsky & Besner, 1993; Borowsky et al., 2013; McCann et al., 1988). The research on bigram frequency effects is far more limited, but bigram frequency effects, whereby increased bigram frequency is related to faster naming RTs, have been attributed to the efficiency of mapping graphemes onto phonemes and thus represent a measure of efficiency of sublexical phonological processing (Borowsky et al., 2013). Given this attribution to phonological processing, we argue that the worse performance for high bigram frequency pseudohomophones represents a form of phonological interference, possibly due to activation of representations in the phonological lexicon via the sublexical grapheme-to-phoneme conversion route.

It is likely the case that both routes are running in parallel (Paap & Noel, 1991), so successful performance in the OLDT simply requires monitoring the orthographic lexicon, which exception words will activate strongly and quickly by utilising word frequency-sensitive connections (Borowsky & Besner, 1993; Borowsky et al., 2013; McCann et al., 1988). The pseudohomophones will not activate the orthographic lexicon as strongly, nor as quickly, so the activation from the bigram frequency-sensitive grapheme-to-phoneme conversion path in the sublexical route will have more opportunity to (incorrectly) intervene. Single-route models will have a difficult time accounting for this double dissociation because they can only make use of one nonsemantic path from orthographic to phonological representations to handle the opposing effects.

Given this double dissociation, our results suggest that both lexical and sublexical processing benefitted from attentional exercise in our studies. This is an exciting extension to the past research on attentional training during reading, as most previous studies have focused on the benefits of attentional training (through video game play) on sublexical processes specifically (Bertoni et al., 2024; Franceschini & Bertoni, 2019). Future studies could use reading tasks that isolate phonetic decoding, such as reading exception words and pseudohomophones aloud (e.g., Kress et al., 2023) or a phonological LDT with pseudohomophone targets and nonword foils (e.g., McCann et al., 1988) to further contrast the impact of short-term attentional exercise on the two reading systems.

In light of the current results, we should also consider the results of the naming task in Kress et al. (2023), where we observed experience with peripheral visual demands was beneficial whereas experience with central visual demands was detrimental to reading aloud performance. These results are somewhat different from those observed in our current experiments. Experiment 1 was consistent with the Kress et al. (2023) results, whereas Experiments 2 and 3 replicated the peripheral but not central A-O exercise effects. One common difference that exists in Kress et al. (2023) and Experiment 1, where the double dissociation between peripheral and central A-O exercise exists, relative to Experiments 2 and 3, is the mixing of peripheral and central visual demands. The A-O exercise in Kress et al. (2023) is provided by the participants' video game experience, and video games include a mix of peripheral and central visual demands. Although the location of stimulus presentation was blocked, Experiment 1 included a combination of peripheral and central A-O blocks to serve as both training and testing measures, so its design is more mixed than Experiments 2 or 3, where only one type of A-O exercise was presented. It could be the case that mixed A-O exercise captures the shared benefits of both types of visual demands, allowing peripheral-exclusive A-O exercise benefits to be observed.

Limitations and future directions

One interesting issue to consider is the case of bigram legality in pseudohomophone construction. Our pseudohomophones were constructed with the intent to use legal bigrams where possible, but in some instances, a very rare (e.g., *uz* in *duz*—found in real words such as *fuzz; wv* in *swawv*—bigram frequency of 0 according to the English Lexicon Project) was used. To address this concern, we repeated our correlational analyses, excluding the exception word/pseudohomophone pairs where the pseudohomophone included a rare bigram, and saw consistent patterns in our correlation (see the online Supplementary Material E). Nevertheless, future studies should consider constructing pseudohomophones and nonwords with systematically varying levels of bigram legality to further assess the differences in orthographic and phonological processing.

In addition, there has been some debate over the consistency of bigram frequency effects (e.g., Schmalz & Mulatti, 2017). However, mixed exception/regular word stimulus lists (such as those used by Schmalz & Mulatti, 2017) are more likely to result in inconsistent bigram frequency effects, because regular words can utilise the dorsal stream grapheme-to-phoneme conversion system (and thus can be affected by bigram frequency), whereas exception words cannot use grapheme-to-phoneme conversion to pronounce the word or make a word/nonword decision (and thus should not be affected by bigram frequency in the same way). Future studies should explicitly test whether exception words and regular words demonstrate different bigram frequency effect patterns, and we may expect that the bigram frequency should have a facilitatory effect on lexical decision performance for regular words, unlike exception words (no effect, as shown here) and pseudohomophones (detrimental effect, as shown here).

The current study focused on measures meant to capture orthographic (word frequency) and phonological processing (bigram frequency). In future studies, it would also be useful to measure the degree of semantic access in this task using a semantic measure such as participant imageability ratings for the base words.

A question raised by a reviewer was whether demographic differences could potentially contribute to the different patterns of results across the three experiments (Experiment 1 participants were the oldest and were not prescreened for education levels, whereas Experiment 2 had the lowest video game experience). However, we consider this unlikely, because the different demographics could counteract each other (e.g., age differences could be balanced by education differences or video game experience). To fully evaluate demographic differences, future studies should prescreen the participants on their demographic factors and baseline performance, and then make sure the two training groups are matched on these factors during the group assignment, or include continuous measures of these as variables in the analysis.

Other paradigm refinements should include manipulations of the exercise duration, such as by comparing single-session designs to multisession designs. Manipulating the exercise duration will determine where the optimal window is to maximise the benefits of A-O exercise while minimising the temporal costs for the individual. For example, reading intervention programmes in schools may have sessions cut short because of limited time and resources (Hall et al., 2022). By identifying the optimal A-O exercise window, a programme could be designed that would better accommodate the limited time and resources and thus be able to reach more individuals who could benefit from such exercise. In comparisons of single-session versus multisession designs, researchers may also be able to bring participants back for delayed post-tests to evaluate the persistency of the benefits gained through large or small A-O exercise (e.g., Pasqualotto et al., 2022).

Conclusion

Our set of experiments manipulated the location of stimuli to test whether A-O exercise may support reading processes. Although Experiment 1 offered some support of our A-O exercise hypothesis, Experiments 2 and 3 raised questions about whether large or small oculomotor shifts matter most in visual attention processes when trying to improve reading ability. Regardless, when the results of these studies are considered together, general learning effects are insufficient to explain the results, and A-O exercise, even if the movements are small, appears to play a role in the reading improvements observed. Future research comparing our small and large movements in A-O exercise to a zero A-O exercise condition will be an important next step in measuring the extent of A-O exercise benefits.

If future studies successfully narrow down an A-O exercise effect, programmes could be developed to specifically target particular reading processes. This will benefit individuals with dyslexia, as they can focus on the exercise that is most relevant to their individual reading deficit. For example, someone with surface dyslexia can focus on the exercise that targets orthographic processing, whereas someone with phonological dyslexia can focus on the exercise associated with phonological processing (Castles & Coltheart, 1993; Franceschini & Bertoni, 2019; McDougall et al., 2005; Sotiropoulos & Hanley, 2017).

Short A-O exercise activities such as a "game-like" version of the task in these studies could provide a boost in reading performance that is worthy of future research. With further study, the A-O exercise demonstrated in our experiments can be refined to be a quick and engaging activity to add to the toolkits of educators who support individuals with reading difficulties.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through an Alexander Graham Bell Canada Graduate Scholarship to S.K., Postdoctoral Fellowship to J.N., and Discovery Grant 18968-2013-25 to the senior author R.B.

ORCID iDs

Shaylyn Kress https://orcid.org/0000-0003-3526-2900
Scott Caron https://orcid.org/0009-0005-4560-7097
Josh Neudorf https://orcid.org/0000-0001-9227-1358
Braedyn Borowsky https://orcid.org/0009-0007-9312-3382
Ron Borowsky https://orcid.org/0000-0003-0002-4021

Data accessibility statement

Data are available on the Open Science Framework at https://doi.org/10.17605/OSF.IO/DGXCJ. Code and experiment materials are available upon reasonable request.

Supplementary material

The supplementary material is available at: qjep.sagepub.com.

Notes

1. Online studies are known for issues with data quality (Douglas et al., 2023; Thomas & Clifford, 2017). We originally recruited 80 participants, and exclusions were made as follows to ensure high data quality:Completed portions of the study twice, compromising their data (two participants).Demographics discrepant with our prescreening criteria (two participants).Reported issues with viewing the instructions properly (one participant).Stratified reaction times, where the reaction time values presented as discrete bands rather than a continuous value, suggesting issues with the keyboard polling rate, or missing response in over half the trials of at least one block, suggesting hardware issues with the keyboard-press detection or participants failing to pay attention (four participants)Unusual error rate or reaction time performance, where error rates were exceedingly high or reaction times exceedingly slow or fast, suggesting participants attempting to exploit the experiment (e.g., always pressing the same button) or failing to follow directions (10 participants)
2. The flankers in the OLDT, although task irrelevant, are intended to help equate the dimensions of the OLDT and HBDT stimuli so that the size and visual complexity of the stimuli is similar.
3. Forty English first-language participants took part in the study, but one was excluded during the data quality check prior to analysis for outlier error rates during the OLDT.

References

Acha, J., Ibaibarriaga, G., Rodríguez, N., & Perea, M. (2024). Lexical and sublexical skills in children's literacy. *Journal of Literacy Research*, *56*, 6–26. https://doi.org/10.1177/1086296X241226476

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01>

Bertoni, S., Andreola, C., Mascheretti, S., Franceschini, S., Ruffino, M., Trezzi, V., Molteni, M., Sali, M. E., Salandi, A., Gaggi, O., Palazzi, C., Gori, S., & Facoetti, A. (2024). Action video games normalise the phonemic awareness in pre-readers at risk for developmental dyslexia. *npj Science of Learning*, *9*(1), 25. https://doi.org/10.1038/s41539-024-00230-0

Bertoni, S., Franceschini, S., Puccio, G., Mancarella, M., Gori, S., & Facoetti, A. (2021). Action video games enhance attentional control and phonological decoding in children with developmental dyslexia. *Brain Sciences*, *11*(2), 171. https://doi.org/10.3390/brainsci11020171

Besner, D., & Borowsky, R. (2006). Postscript: Plaut and Booth's (2006) new simulations–what have we learned? *Psychological Review*, *113*(1), 194–195. https://doi.org/10.1037/0033-295X.113.1.194

Boden, C., & Giaschi, D. (2007). M-stream deficits and reading-related visual processes in developmental dyslexia. *Psychological Bulletin*, *133*(2), 346–366. https://doi.org/10.1037/0033-2909.133.2.346

Boros, M., Anton, J.-L., Pech-Georgel, C., Grainger, J., Szwed, M., & Ziegler, J. C. (2016). Orthographic processing deficits in developmental dyslexia: Beyond the ventral visual stream. *NeuroImage*, *128*, 316–327. https://doi.org/10.1016/j.neuroimage.2016.01.014

Borowsky, R., & Besner, D. (1993). Visual word recognition: A multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 813–840.

Borowsky, R., & Besner, D. (2006). Parallel distributed processing and lexical-semantic effects in visual word recognition: Are a few stages necessary? *Psychological Review*, *113*(1), 181–193. https://doi.org/10.1037/0033-295X.113.1.181

Borowsky, R., Cummine, J., Owen, W. J., Friesen, C. K., Shih, F., & Sarty, G. E. (2006). FMRI of ventral and dorsal processing streams in basic reading processes: Insular sensitivity to phonology. *Brain Topography*, *18*(4), 233–239. https://doi.org/10.1007/s10548-006-0001-2

Borowsky, R., Esopenko, C., Gould, L., Kuhlmann, N., Sarty, G., & Cummine, J. (2013). Localisation of function for noun and verb reading: Converging evidence for shared processing from fMRI activation and reaction time. *Language and Cognitive Processes*, *28*(6), 789–809. https://doi.org/10.1080/01690965.2012.665466

Brochu, P., O'Grady, K., Scerbina, T., & Tao, Y. (2018). *PIRLS/ePIRLS 2016 Canada in context: Canadian results from the progress in international reading literacy study*. Council of Ministers of Education, Canada. https://cmec.

ca/Publications/Lists/Publications/Attachments/385/PIRLS2016-Report-EN.pdf

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, *47*(2), 149–180. https://doi.org/10.1016/0010-0277(93)90003-E

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Council of Ministers of Education Canada & Statistics Canada. (2013). *Skills in Canada: First results from the Programme for the International Assessment of Adult Competencies (PIAAC)*. Statistics Canada/Statistique Canada.

Cummine, J., Dai, W., Borowsky, R., Gould, L., Rollans, C., & Boliek, C. (2015). Investigating the ventral-lexical, dorsal-sublexical model of basic reading processes using diffusion tensor imaging. *Brain Structure and Function*, *220*(1), 445–455. https://doi.org/10.1007/s00429-013-0666-8

Diarra, M., Zendel, B. R., Benady-Chorney, J., Blanchette, C.-A., Lepore, F., Peretz, I., Belleville, S., & West, G. L. (2019). Playing Super Mario increases oculomotor inhibition and frontal eye field grey matter in older adults. *Experimental Brain Research*, *237*(3), 723–733. https://doi.org/10.1007/s00221-018-5453-6

Douglas, B. D., Ewell, P. J., & Brauer, M. (2023). Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLOS ONE*, *18*(3), Article e0279720. https://doi.org/10.1371/journal.pone.0279720

Dye, M. W. G., Green, C. S., & Bavelier, D. (2009). The development of attention skills in action video game players. *Neuropsychologia*, *47*(8–9), 1780–1789. https://doi.org/10.1016/j.neuropsychologia.2009.02.002

Ekstrand, C., Gould, L., Mickleborough, M., Lorentz, E., & Borowsky, R. (2016). When words and space collide: Spatial attention interacts with lexical access during word recognition. *Visual Cognition*, *24*(3), 284–291. https://doi.org/10.1080/13506285.2016.1236867

Ekstrand, C., Neudorf, J., Gould, L., Mickleborough, M., & Borowsky, R. (2019). Where words and space collide: The overlapping neural activation of lexical and sublexical reading with voluntary and reflexive spatial attention. *Brain Research*, *1706*, 1–12. https://doi.org/10.1016/j.brainres.2018.10.022

Ekstrand, C., Neudorf, J., Kress, S., & Borowsky, R. (2019). How words and space collide: Lexical and sublexical reading are reliant on separable reflexive and voluntary attention regions in hybrid tasks. *Cortex*, *121*, 104–116. https://doi.org/10.1016/j.cortex.2019.08.006

Facoetti, A., Paganoni, P., Turatto, M., Marzola, V., & Mascetti, G. (2000). Visual-spatial attention in developmental dyslexia. *Cortex*, *36*(1), 109–123. https://doi.org/10.1016/S0010-9452(08)70840-2

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14*(3), 340–347. https://doi.org/10.1162/089892902317361886

Foorman, B. R., Herrera, S., & Dombek, J. (2018). The relative impact of aligning Tier 2 intervention materials with classroom core reading materials in grades K–2. *The Elementary School Journal*, *118*(3), 477–504. https://doi.org/10.1086/696021

Franceschini, S., & Bertoni, S. (2019). Improving action video games abilities increases the phonological decoding speed and phonological short-term memory in children with developmental dyslexia. *Neuropsychologia*, *130*, 100–106. https://doi.org/10.1016/j.neuropsychologia.2018.10.023

Franceschini, S., Trevisan, P., Ronconi, L., Bertoni, S., Colmar, S., Double, K., Facoetti, A., & Gori, S. (2017). Action video games improve reading abilities and visual-to-auditory attentional shifting in English-speaking children with dyslexia. *Scientific Reports*, *7*(1), 5863. https://doi.org/10.1038/s41598-017-05826-8

Gabrieli, J. D. E., & Norton, E. S. (2012). Reading abilities: Importance of visual-spatial attention. *Current Biology*, *22*(9), R298–R299. https://doi.org/10.1016/j.cub.2012.03.041

Gori, S., Cecchini, P., Bigoni, A., Molteni, M., & Facoetti, A. (2014). Magnocellular-dorsal pathway and sub-lexical route in developmental dyslexia. *Frontiers in Human Neuroscience*, *8*, Article 460. https://doi.org/10.3389/fnhum.2014.00460

Gori, S., Seitz, A. R., Ronconi, L., Franceschini, S., & Facoetti, A. (2016). Multiple causal links between magnocellular–dorsal pathway deficit and developmental dyslexia. *Cerebral Cortex*, *26*(11), 4356–4369. https://doi.org/10.1093/cercor/bhv206

Hall, C., Dahl-Leonard, K., & Cannon, G. (2022). Observing two reading intervention programs for students with dyslexia. *Exceptionality*, *30*(2), 109–125. https://doi.org/10.1080/09362835.2021.1938067

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*(3), 491–528. https://doi.org/10.1037/0033-295X.106.3.491

Holmes, J., Woolgar, F., Hampshire, A., & Gathercole, S. E. (2019). Are working memory training effects paradigm-specific? *Frontiers in Psychology*, *10*, Article 1103. https://doi.org/10.3389/fpsyg.2019.01103

Kowalczyk, N., Shi, F., Magnuski, M., Skorko, M., Dobrowolski, P., Kossowski, B., Marchewka, A., Bielecki, M., Kossut, M., & Brzezicka, A. (2018). Real-time strategy video game experience and structural connectivity—A diffusion tensor imaging study. *Human Brain Mapping*, *39*(9), 3742–3758. https://doi.org/10.1002/hbm.24208

Kress, S., Neudorf, J., Borowsky, B., & Borowsky, R. (2023). What's in a game: Video game visual-spatial demand location exhibits a double dissociation with reading speed. *Acta Psychologica*, *232*, 103822. https://doi.org/10.1016/j.actpsy.2022.103822

Kress, S., Neudorf, J., Ekstrand, C., & Borowsky, R. (2021). Unmasking the effects of orthography, semantics, and phonology on 2AFC visual word perceptual identification. *Visual Cognition*, *29*, 657–678. https://doi.org/10.1080/13506285.2021.1989099

Li, X., & Pollatsek, A. (2020). An integrated model of word processing and eye-movement control during Chinese reading. *Psychological Review*, *127*(6), 1139–1162. https://doi.org/10.1037/rev0000248

Masson, M. E. J., & Borowsky, R. (1998). More than meets the eye: Context effects in word identification. *Memory & Cognition*, *26*(6), 1245–1269. https://doi.org/10.3758/BF03201198

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *57*(3), 203–220. https://doi.org/10.1037/h0087426

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McCann, R. S., Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word-frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 693–706.

McDougall, P., Borowsky, R., MacKinnon, G. E., & Hymel, S. (2005). Process dissociation of sight vocabulary and phonetic decoding in reading: A new perspective on surface and phonological dyslexias. *Brain and Language*, *92*(2), 185–203. https://doi.org/10.1016/j.bandl.2004.06.003

Minissi, M. E., Antzaka, A., Mancini, S., & Lallier, M. (2024). Can playing video games enhance reading skills through more efficient serial visual search mechanisms? Insights from an eye tracking study. *Language, Cognition and Neuroscience*. https://doi.org/10.1080/23273798.2024.2411696

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 667–682.

Morys-Carter, W. L. (2020). *ScreenScale* [Python]. https://doi.org/10.17605/OSF.IO/8FHQK

Neudorf, J., Ekstrand, C., Kress, S., & Borowsky, R. (2019). FMRI of shared-stream priming of lexical identification by object semantics along the ventral visual processing stream. *Neuropsychologia*, *133*, 107185. https://doi.org/10.1016/j.neuropsychologia.2019.107185

Neudorf, J., Ekstrand, C., Kress, S., Neufeldt, A., & Borowsky, R. (2019). Interactions of reading and semantics along the ventral visual processing stream. *Visual Cognition*, *27*(1), 21–37. https://doi.org/10.1080/13506285.2019.1577319

Nouchi, R., Taki, Y., Takeuchi, H., Hashizume, H., Akitsuki, Y., Shigemune, Y., Sekiguchi, A., Kotozaki, Y., Tsukiura, T., Yomogida, Y., & Kawashima, R. (2012). Brain training game improves executive functions and processing speed in the elderly: A randomized controlled trial. *PLOS ONE*, *7*(1), Article e29676. https://doi.org/10.1371/journal.pone.0029676

Owen, W. J., & Borowsky, R. (2003). Examining the interactivity of lexical orthographic and phonological processing. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *57*(4), 290–303. https://doi.org/10.1037/h0087432

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research*, *53*(1), 13–24. https://doi.org/10.1007/BF00867328

Partanen, M., Siegel, L. S., & Giaschi, D. E. (2019). Longitudinal outcomes of an individualized and intensive reading intervention for third grade students. *Dyslexia*, *25*, 227–245. https://doi.org/10.1002/dys.1616

Pasqualotto, A., Altarelli, I., De Angeli, A., Menestrina, Z., Bavelier, D., & Venuti, P. (2022). Enhancing reading skills through a video game mixing action mechanics and cognitive training. *Nature Human Behaviour*, *6*, 545–554. https://doi.org/10.1038/s41562-021-01254-x

Patterson, K., & Hodges, J. R. (1992). Deterioration of word meaning: Implications for reading. *Neuropsychologia*, *30*(12), 1025–1040.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pfost, M., Dörfler, T., & Artelt, C. (2013). Students' extracurricular reading behavior and the development of vocabulary and reading comprehension. *Learning and Individual Differences*, *26*, 89–102. https://doi.org/10.1016/j.lindif.2013.04.008

Plaut, D. C., & Booth, J. R. (2006). More modeling but still no stages: Reply to Borowsky and Besner. *Psychological Review*, *113*(1), 196–200. https://doi.org/10.1037/0033-295X.113.1.196

PsychoPy (Version 2022.2.4) [Computer software]. (2022). Open Science Tools Ltd. https://psychopy.org/

PsychoPy (Version 2023.1.3) [Computer software]. (2023). Open Science Tools Ltd. https://psychopy.org/

Puccio, G., Gazzi, G., Bertoni, S., Franceschini, S., Mancarella, M., Gori, S., Ronconi, L., & Facoetti, A. (2023). Action video games training in children with developmental dyslexia: A meta-analysis. *International Journal of Human–Computer Interaction*, *40*, 6829–6844. https://doi.org/10.1080/10447318.2023.2267297

R Core Team. (2022). *R: A language and environment for statistical computing* (Version 4.2.2) [Computer software]. www.R-project.org

Schmalz, X., & Mulatti, C. (2017). Busting a myth with the Bayes Factor: Effects of letter bigram frequency in visual lexical decision do not reflect reading processes. *The Mental Lexicon*, *12*(2), 263–282. https://doi.org/10.1075/ml.17009.sch

Sella, F., Tressoldi, P., Lucangeli, D., & Zorzi, M. (2016). Training numerical skills with the adaptive videogame "The Number Race": A randomized controlled trial on preschoolers. *Trends in Neuroscience and Education*, *5*(1), 20–29. https://doi.org/10.1016/j.tine.2016.02.002

Sotiropoulos, A., & Hanley, J. R. (2017). Developmental surface and phonological dyslexia in both Greek and English. *Cognition*, *168*, 205–216. https://doi.org/10.1016/j.cognition.2017.06.024

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://psycnet.apa.org/doi/10.1598/RRQ.21.4.1

Taran, N., Farah, R., DiFrancesco, M., Altaye, M., Vannest, J., Holland, S., Rosch, K., Schlaggar, B. L., & Horowitz-Kraus, T. (2022). The role of visual attention in dyslexia: Behavioral and neurobiological evidence. *Human Brain Mapping*, *43*(5), 1720–1737. https://doi.org/10.1002/hbm.25753

Tatler, B. W., Brockmole, J. R., & Carpenter, R. H. S. (2017). LATEST: A model of saccadic decisions in space and time. *Psychological Review*, *124*(3), 267–300. https://doi.org/10.1037/rev0000054

Thomas, K. A., & Clifford, S. (2017). Validity and Mechanical Turk: An assessment of exclusion methods and interactive experiments. *Computers in Human Behavior*, *77*, 184–197. https://doi.org/10.1016/j.chb.2017.08.038

Vidyasagar, T. R., & Pammer, K. (2010). Dyslexia: A deficit in visuo-spatial attention, not in phonological processing. *Trends in Cognitive Sciences*, *14*(2), 57–63. https://doi.org/10.1016/j.tics.2009.12.003

West, G. L., Al-Aidroos, N., & Pratt, J. (2013). Action video game experience affects oculomotor performance. *Acta Psychologica*, *142*(1), 38–42. https://doi.org/10.1016/j.actpsy.2011.08.005

Wingerak, S., Neudorf, J., Gould, L., & Borowsky, R. (2017). On the dissociation between reaction time and response duration as a function of lexical and sublexical reading: An examination of phonetic decoding and computational models. *Visual Cognition*, *25*(9–10), 913–927. https://doi.org/10.1080/13506285.2017.1352640