Journal of Neurolinguistics 51 (2019) 309–323

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

# Concrete and abstract word processing in deep dyslexia

Simritpal Kaur Malhi, Tara Lynn McAuley, Brette Lansue, Lori Buchanan*

*Department of Psychology, University of Windsor, 401 Sunset Avenue, Windsor, Ontario, N9B 3P4, Canada*

ARTICLE INFO

*Keywords:*
Deep dyslexia
Failure of inhibition theory (FIT)
Concrete word processing
Abstract word processing

ABSTRACT

The purpose of this case study was to test the failure of inhibition theory of deep dyslexia (FIT; Buchanan, McEwen, Westbury, & Libben, 2003) with concrete and abstract words. FIT proposes that in deep dyslexia, errors to abstract words are the result of an impairment in phonological output lexicon selection rather than a semantic deficit for abstract words. FIT also proposes a dissociation between explicit phonological lexicon production (can be compromised) and implicit access of representations (is intact). With such assumptions it follows that in phonologically implicit tasks where controls demonstrate either concreteness or abstractness effects, a participant with deep dyslexia would similarly show concreteness or abstractness effects. However, for explicit tasks where production is involved, a participant with deep dyslexia would only show concreteness effects due to difficulty with abstract word production, indicative of their difficulty with phonological output lexicon selection which is more compromised for abstract words because semantic content cannot guide the selection. Experiments 1–3 used phonologically implicit tasks (i.e., concrete categorization task, semantic relatedness task, and iconicity judgment task) and Experiment 4 used an explicit task (i.e., oral word-reading task). The results supported the hypotheses and are consistent with FIT as an explanation for the locus of impairment in deep dyslexia.

## 1. Introduction

Deep dyslexia is an acquired reading disorder that results from damage to language areas in the left hemisphere of the brain (Marshall & Newcombe, 1973). Unlike other acquired reading disorders (e.g., surface dyslexia, phonological dyslexia, and pure alexia), the neural correlates of deep dyslexia are less conclusive, and the characteristics of deep dyslexia often manifest from diffuse left hemisphere lesions (Ripamonti et al., 2014). The constellation of symptoms that are characteristic of deep dyslexia include semantic errors (e.g., *tide* for *water*), visual errors (e.g., *month* for *mouth*), phonological errors (e.g., *smile* for *style* but see Buchanan, Hildebrandt, & MacKinnon, 1994 for arguments for considering these as visual errors), and morphological errors (e.g., *showing* for *shown*) during single-word oral reading, as well as an inability to read pseudowords (e.g., *JUHT*) aloud (see Coltheart, 1980a for a review of types of errors in deep dyslexia). Another characteristic of deep dyslexia reported in the literature is the concreteness effect (i.e., a processing advantage for concrete words) whereby concrete words (e.g., *table*) are read aloud more accurately than abstract words (e.g., *justice*; Coltheart, Patterson, & Marshall, 1980; Newton & Barry, 1997).

The concreteness effect not only occurs in data from individuals with deep dyslexia, but also in controls (Paivio, 1991). Major theories explaining the concreteness effect in unimpaired readers include the dual-coding theory (Paivio, 1991), which argues that concrete words have a processing advantage because they activate both the linguistic and imagistic systems, whereas abstract words

* Corresponding author.
*E-mail address:* buchanan@uwindsor.ca (L. Buchanan).

https://doi.org/10.1016/j.jneuroling.2018.11.001
Received 20 October 2017; Received in revised form 26 October 2018; Accepted 3 November 2018
Available online 23 November 2018
0911-6044/ 

only activate the linguistic system. Another explanation for the concreteness effect is the context-availability theory (Schwanenflugel, Harnishfeger, & Stowe, 1988; Schwanenflugel & Shoben, 1983; Schwanenflugel & Stowe, 1989), which argues that concrete words are strongly associated with a few contexts, whereas abstract words are weakly associated with many contexts.

The concreteness effect is, however, exaggerated in those with deep dyslexia such that there can be a complete inability to read abstract words. There is some evidence to suggest that this exaggerated concreteness effect is also reflected in neural activation differences in controls and persons with aphasia (Sandberg & Kiran, 2014). Various theories explaining the concreteness effect in deep dyslexia have been proposed. According to Coltheart's (1980b) right hemisphere hypothesis, the left hemisphere subserves abstract word reading. Readers with deep dyslexia have damage to the left hemisphere, and thus, rely heavily on the right hemisphere, resulting in difficulty with abstract words. Morton and Patterson (1980) propose a dual-route model in which deep dyslexia results from multiple loci of damage. In this model, reading occurs via the semantic route; however, the semantics for abstract words are impaired. Similarly, Plaut and Shallice (1993) in their connectionist model propose an advantage for reading concrete words because concrete words have more semantic features than abstract words. Further, assuming spreading activation (e.g., McClelland, 1987), Crutch and Warrington's (2005) different representational frameworks model proposes that concrete words are represented in a categorical framework (i.e., based on semantic similarity) and abstract words are represented mainly by semantic association (i.e., linguistics contexts). This theory maintains that concrete words share more representations with other similar words (e.g., *cow–sheep*) than with other associated words (e.g., *cow–barn*) whereas abstract words share more representations with other associated words (e.g., *robbery–punishment*) than with other similar words (e.g., *robbery–theft*). As a result, readers with deep dyslexia produce more associative errors than semantically similar errors in response to abstract target words and more semantically similar errors than associative errors in response to concrete target words (Crutch, 2006).

These aforementioned theories suggest a semantic deficit for abstract words (and a semantic advantage for concrete words) when explaining the concreteness effect in deep dyslexia. That is, abstract words are read aloud inaccurately because an impaired semantic system fails to process the meaning of these words. In contrast, the failure of inhibition theory (FIT; Buchanan, McEwen, Westbury, & Libben, 2003) centers on the claim that semantic processing in deep dyslexia remains intact and that the dysfunction occurs at the level of the phonological output lexicon due to reduced inhibitory processes. This theory, also assuming spreading activation (e.g., McClelland, 1987), proposes that when presented with a target word, that representation as well as other neighbouring representations in the orthographic and semantic systems are activated. For tasks requiring explicit output via reading aloud, activation spreads to the phonological output lexicon in which representations of phonological neighbours of the target are activated. In unimpaired reading aloud, inhibitory processes decrease the activation of orthographic, semantic, and phonological representations so that the target is the most highly activated representation. FIT proposes that in a deep dyslexic reader, these inhibitory processes are slowed or reduced in the phonological output lexicon. This deficit contributes to reading errors, as spurious activations are not successfully inhibited, and neighbours are selected in place of the target. Although FIT would predict other types of reading errors (e.g., failure to inhibit orthographic representations leads to visual reading errors), semantic reading errors in particular characterize deep dyslexic readers due to the failure to inhibit neighbouring semantic representations at the level of the phonological output lexicon. In this way, FIT does not modify the assumption of spreading activation models; rather, it identifies the source of error when reading aloud. That is, the deficit is only obvious when explicit demands of the task require the phonological output lexicon and reveal the failure of inhibition, whereas tasks that rely on implicit access bypass the phonological output lexicon and show a typical pattern of performance (such as a disadvantage for pseudohomophone rejection over pronounceable nonwords in Buchanan et al., 1994). See Fig. 1 for a simple schematic diagram that illustrates FIT.

Therefore, the impairment is at the level of phonological output lexicon selection as opposed to the level of semantics. Buchanan et al. (2003) relate this dissociation between semantics and phonological output lexicon selection to the dissociation between lexical representations and lexical access (e.g., Warrington & Shallice, 1979). According to this model, the errors associated with deep dyslexia arise from a failure to inhibit competing lexical entries in the phonological output lexicon. Such failure results in compromised explicit access and phonological lexicon production but intact implicit access of semantic representations.

Direct and indirect support for FIT comes from various case studies. For example, Buchanan et al. (2003) directly tested FIT and reported a deep dyslexic reader accessing semantic information about both words and pseudohomophones while being unable to produce them. Failure of inhibition is only problematic when a target must be selected from the phonological output lexicon. On tasks where the phonological output lexicon is not required, the performance of those with deep dyslexia would not be expected to differ from that of controls. For example, Colangelo and Buchanan (2005) reported the performance of a deep dyslexic reader who had a specific deficit for ambiguous words during an explicit task (i.e., naming task) but an advantage for the same ambiguous words in an implicit task (i.e., lexical decision task). Colangelo and Buchanan (2006) also found that a deep dyslexic reader was able to explicitly demonstrate access to semantic information regarding associative relationships (i.e., which word is most associated with all other words on the list) when phonological lexicon production was not required.

In addition to intact access to semantic representations, studies supporting FIT have also found intact implicit access of phonological representations (Buchanan et al., 1994; Buchanan, Hildebrandt, & MacKinnon, 1996, 1999). For example, Colangelo and Buchanan (2007) found that a deep dyslexic reader was able to rely on assembled phonology and use semantic information from pseudohomophones to explicitly discriminate semantic information, all the while generating errors during phonological lexicon production. Similarly, while FIT was not directly tested, Marelli, Aggujaro, Molteni, and Luzzatti (2012) concluded that the basis of impairment in a deep dyslexic reader was at the retrieval of a phonological output lexicon representation. Their conclusion was derived from data from an individual with deep dyslexia on both implicit and explicit tasks, including tasks of visual and oral lexical decisions, word comprehension, semantic judgments, picture naming, and reading aloud words from different grammatical classes. Collectively, performance suggested that visual and phonological input lexicons and semantic processing were intact, but an

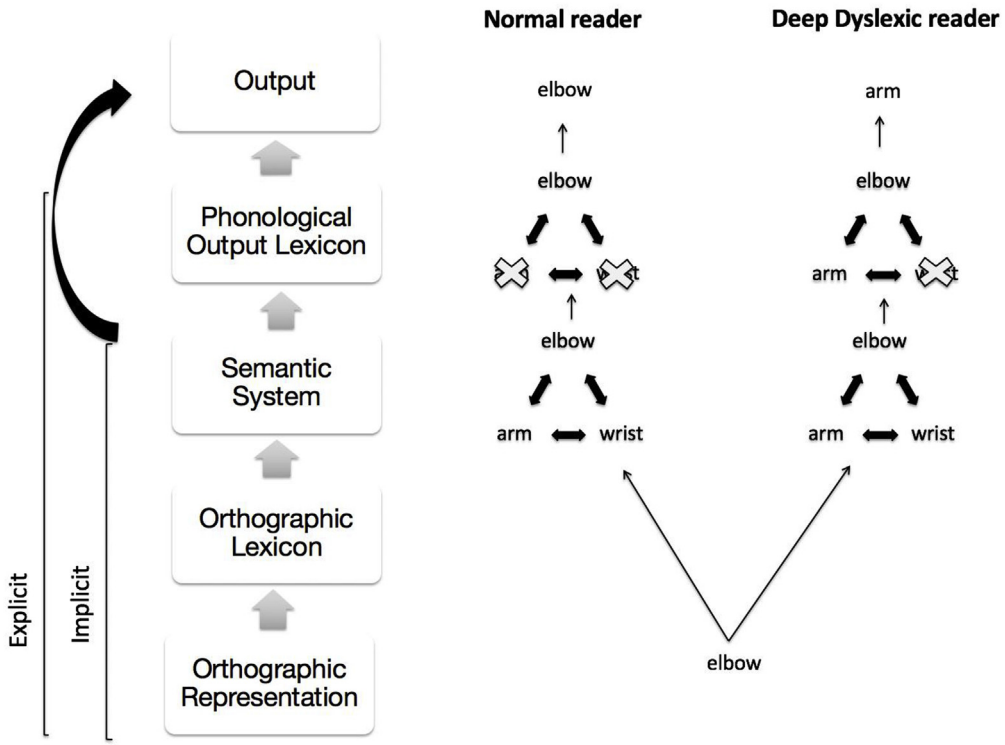


**Fig. 1.** A simple schematic diagram of FIT and examples of explicit access when reading aloud in a control and deep dyslexic reader. Explicit output relies on the phonological output lexicon, whereas implicit output bypasses it. In this example, both readers are presented with the written word "elbow". An intact reader accurately inhibits and selects the appropriate candidate word to produce ("elbow"). A deep dyslexic reader fails to inhibit semantically related words at the level of the phonological output lexicon and produces a semantic reading error ("wrist"). Note that only spreading activation in the semantic system is illustrated; orthographic and phonological neighbour representations were not included for simplicity.

impairment was notable at the level of lexical retrieval from the phonological output lexicon. Additionally, as FIT proposes that spurious phonological output lexicon activations are incorrectly selected in deep dyslexia, support for FIT comes from studies demonstrating that supplying the initial phoneme of a target word improves oral reading in those with deep dyslexia (Buchanan, Kiss, & Burgess, 2000; Katz & Lanzoni, 1997), whereas providing an inaccurate phonemic cue results in poorer reading (Katz & Lanzoni, 1997).

Further, while semantic context facilitates performance in controls, FIT would predict that semantic context increases semantic errors in those with deep dyslexia as more neighbours are activated and thus more candidates are created for the phonological output lexicon. In support of this, Colangelo, Westbury, and Buchanan (2004) found more semantic errors when a deep dyslexic reader was asked to read semantically blocked words (i.e., words grouped by semantic categories, e.g., *table*, *sit*, *seat*, *couch*) than when they were asked to read those words presented in a random order. Similarly, Bose, Colangelo, and Buchanan (2011) found more errors when a deep dyslexic reader read semantically blocked words only when the words were phonetically complex. However, during a repetition task, while more errors were made for phonetically complex words than for phonetically simple words, the deep dyslexic reader's performance was unaffected by semantic blocking. These researchers theorized that processing from orthographic input through semantic activation to speech motor movement is a cascaded process. Therefore, the lack of sensitivity to semantic blocking in the repetition task may have been due to a stronger activation in the phonological output lexicon, eliminating the need for feedback from semantics.

While FIT has not been directly tested with concrete and abstract words, FIT would give rise to the prediction that for phonologically implicit tasks where controls show either a concreteness or an abstractness effect (i.e., a processing advantage for abstract words), those with deep dyslexia would follow the same pattern. However, for explicit tasks in which the phonological output lexicon is required, those with deep dyslexia would only show a concreteness effect due to difficulty with abstract word production. Abstract words are more affected than concrete words with respect to production because the semantic content of abstract words guides their selection less. Concrete words tend to activate few and firm candidates whereas abstract words activate multiple and not firm candidates (e.g., *chair* may activate *desk*, whereas *freedom* may activate *liberty*, *justice*, etc.). There are fewer constraints on abstract words because they have more associates (i.e., semantic neighbours) leading to increased opportunities for errors in a compromised phonological output lexicon. There is a difference between concrete and abstract words at the level of the phonological output lexicon, but only in so far as there is a difference at the semantic level such that more semantic items are going forward for abstract words. Prior studies on concrete and abstract word processing in the deep dyslexia literature appear to support such predictions. For

example, Boumaraf and Macoir (2016) in their case study of an Arabic-speaking patient with deep dyslexia found no significant differences in the comprehension between high imageability and low imageability words in an auditory synonym judgment task and only marginally significant differences in comprehension between high imageability and low imageability words in a visual synonym judgment task. However, they found a significant difference in reading aloud high imageability and low imageability words, with better accuracy for the former. Moreover, Newton and Barry (1997) examined the comprehension of abstract words in a deep dyslexic reader and found that while she was better able to read aloud concrete words compared to abstract words, she had no significant impairment with respect to her comprehension of abstract high-frequency words. These researchers concluded that the semantic representations of abstract words are intact in deep dyslexia. They proposed the NICE (Normal Isolated Centrally Expressed semantics) model in which the semantics of concrete words have more specificity such that they uniquely activate the target word in one's lexicon for phonological output. In contrast, abstract words have less specificity and greater spreading activation to other competing words, which in turn results in incorrect selection and semantic errors made by those with deep dyslexia. These findings are consistent with FIT, and Newton and Barry (1997) even suggested that reading in deep dyslexia may reveal more about spoken word production than about word comprehension, and as such, deep dyslexia may be more aptly labeled as "deep production." Others have argued that the concreteness effect in deep dyslexia is the result of both a semantic deficit and a deficit in phonological output lexicon retrieval (Barry & Gerhand, 2003; Riley & Thompson, 2010). This suggestion is in contrast to Buchanan et al.'s (2003) claim that semantic processing remains intact and is not entirely consistent with the evidence described above. The assumptions of the FIT model *vis a vis* concreteness effects is therefore the subject of this case study.

The goal of this paper is to directly test FIT using both concrete and abstract words. Both phonologically implicit tasks (i.e., concrete categorization task, semantic relatedness task, and an iconicity judgment task) and an explicit task (i.e., oral word-reading task) will be used. It is hypothesized that in the phonologically implicit tasks, performance of the participant with deep dyslexia will map onto the performance of controls (i.e., unimpaired readers) both from this study and those who have previously completed these tasks with similar or identical stimuli (Friederici, Opitz, & von Cramon, 2000; Malhi & Buchanan, 2018), and thus not reflect a semantic deficit for abstract words. Friederici et al. (2000) compared concrete and abstract nouns and function words in both a semantic task (i.e., is this a concrete or an abstract word) and a syntactic task (i.e., is this a noun or a function word) with controls. Results showed an interaction between concreteness and word class in both tasks. Specifically, controls showed a concreteness effect for nouns, where they were faster to categorize concrete nouns. Malhi and Buchanan (2018) studied concrete (e.g., *nose—tongue*) and abstract (e.g., *gain—loss*) word pairs in both a semantic relatedness task (i.e., are these word pairs related or unrelated) and an iconicity judgment task (i.e., are these word pairs iconic or reverse-iconic) with controls. In addition to word pairs being either concrete or abstract, word pairs were either iconic (e.g., *nose* above *tongue*; *gain* above *loss*) or reverse-iconic (e.g., *tongue* above *nose*; *loss* above *gain*), and either close (e.g., *nose* is the ninth neighbour of *tongue*, and *tongue* is the 22nd neighbour of *nose*) or distant (e.g., *gain* is the 305$^{th}$ neighbour of *loss*, and *loss* is the 394$^{th}$ neighbour of *gain*) semantic neighbours based on semantic neighbourhood distance values obtained from WINDSORS (Durda & Buchanan, 2008). Results showed that in the semantic relatedness task, close rather than distant semantic neighbours facilitated processing, and in the iconicity judgment task, iconic rather than reverse-iconic presentation facilitated processing. Across both tasks, controls showed an abstractness effect, which was interpreted in the context of the nature of the tasks. Given that these tasks produced an abstractness effect in controls, they provide an opportunity to study the possibility of intact semantic processing of abstract words in deep dyslexia. It is hypothesized that while the participant with deep dyslexia will be slower and make more errors overall, a similar pattern as controls is expected, whether that is concreteness (Friederici et al., 2000) or abstractness (Malhi & Buchanan, 2018). It is important to note that the very presence of concreteness or abstractness effects reflects semantic processing.

However, in the phonologically explicit task that relies on the phonological output lexicon, it is hypothesized that the participant with deep dyslexia will demonstrate only a concreteness effect due to difficulty with abstract word production, and thus reflect a phonological output lexicon selection deficit for abstract words. For the explicit task, performance of the participant with deep dyslexia is expected to diverge from the performance of controls, such that readers with deep dyslexia would show a concreteness effect due to difficulty with abstract word production, whereas controls would not have difficulty with abstract word production. See Table 1 for a summary of hypotheses.

**Table 1**
Summary of hypotheses for all experiments.

| Experiment | Type of Task | Performance of Controls | Hypothesis for Deep Dyslexic Reader |
|---|---|---|---|
| Experiment 1: Concrete categorization task | Implicit | Concreteness effect | Concreteness effect |
| Experiment 2: Semantic relatedness task | Implicit | Abstractness effect | Abstractness effect |
| Experiment 3: Iconicity judgment task | Implicit | Abstractness effect | Abstractness effect |
| Experiment 4: Oral word-reading task | Explicit | No difference between concrete and abstract word production | Difficulty with abstract word production |

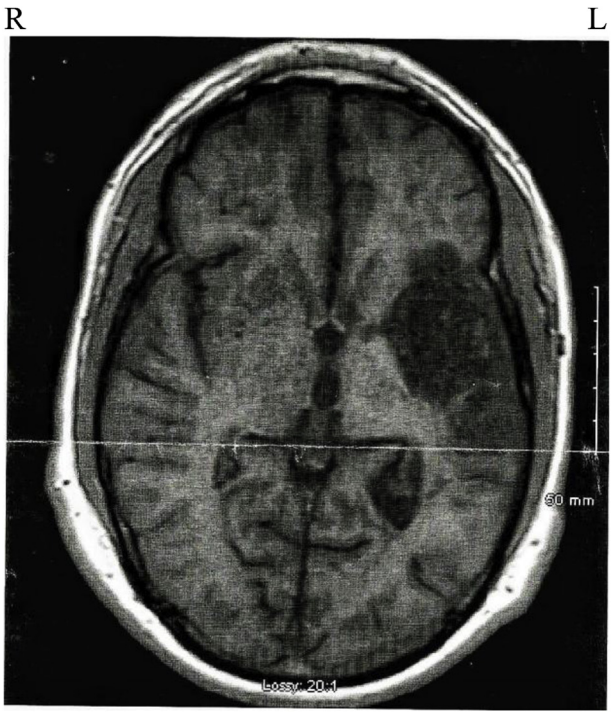


**Fig. 2.** Computed Tomography (CT) scan of GL following a large, left MCA stroke.

## 2. Method

### *2.1. Participant*

#### *2.1.1. Demographics and stroke details*

GL is a 35-year old male with 17 years of formal education. Approximately 8 years ago, GL was diagnosed with a cerebrovascular accident (CVA) in the left middle cerebral artery (MCA). Computed Tomography (CT) scans confirmed that the left MCA stroke (see Fig. 2) was secondary to a left internal carotid artery dissection (see Fig. 3). The stroke resulted in hemiparesis on the right side of his body and aphasia. Following the stroke, GL experienced complications that included grand-mal seizures and post-stroke depression. GL has been reasonably healthy since then and has had speech therapy sporadically.

Controls ($n$ = 16, four males, 12 females; *Mage* = 22.87 years, age range: 18–44 years) for Experiment 1 were undergraduate psychology students who learned English as their first language. For Experiments 2 and 3, control data from Malhi and Buchanan (2018) were used.

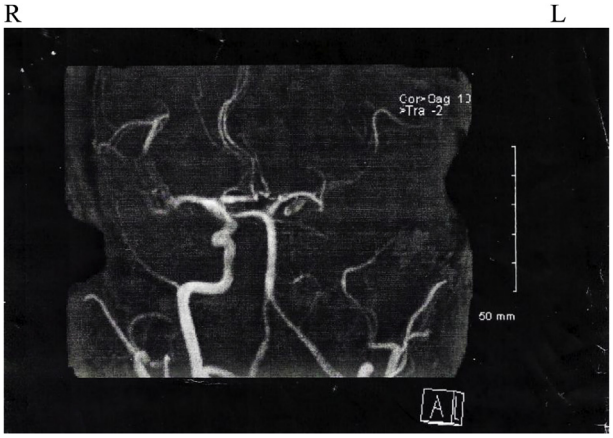


**Fig. 3.** Magnetic Resonance Angiogram (MRA) confirming the left internal carotid artery dissection.

### 2.1.2. Speech-language assessment pre- and post-treatment

During an initial speech, language, and cognitive assessment, GL met criteria for moderate to severe deficits in auditory comprehension, reading comprehension, verbal expression, and written expression. These results were based on responses to a standard aphasia protocol given by a speech language pathologist in a rehabilitation setting, which included assessing consistency in yes/no responding, following commands, reading simple sentences, repetition, picture-word discrimination, matching pictures to words, naming objects, responsive naming, sentence production, and written word production. The swallowing assessment, speech/voice assessment, hearing assessment, and cognitive assessment (attention and concentration were only assessed at this time) were all determined to be within functional limits. GL was recommended speech, language, and cognitive-communication therapy (i.e., since all aspects of cognition relevant to communication could not be tested due to dependency on language for task completion). Upon discharge from rehabilitation, GL was re-assessed on the same standard aphasia protocol, and improvements were found in auditory comprehension, reading comprehension, verbal expression, and written expression. To elaborate, GL was determined to be within normal functional limits for responses to simple yes/no questions and reading letters and numbers. He was also within normal functional limits on tests of attention and concentration, immediate memory, and short-term memory. He demonstrated a mild deficit in following one step commands, reading single words, repeating short strings of words, and picture-word discrimination. At this time, GL demonstrated a moderate to severe deficit in understanding complex yes/no questions, following two or three step commands, understanding paragraphs, reading simple sentences and paragraphs, naming objects, producing sentences, and writing words. Overall, at the time of the post-assessment, GL continued to demonstrate receptive and expressive language difficulties, perseverate on words and thoughts, and produce phonemic and verbal paraphasias.

### 2.1.3. Results of word-reading tasks

Similar to other case studies on deep dyslexic readers (e.g., Buchanan et al., 2003; Colangelo, Stephenson, Westbury, & Buchanan, 2003), GL was tested on word reading tasks to determine whether he was a deep dyslexic reader. On these tasks, GL produced multiple reading errors involving a variety of different word classes. See Table 2 for examples of reading errors made on these tasks. Based on this error profile, and in particular, the presence of semantic errors that characterize deep dyslexic readers, GL meets criteria for deep dyslexia. Morphological reading errors also characterize deep dyslexic readers (Coltheart, 1981), and in English they are predominant on the right side of a target word. This might lead one to suspect a failure to follow instruction or a visual field cut but GL has performed numerous experimental tasks in our lab, both linguistic and non-linguistic (e.g., picture naming, matching tasks), and he has been able to follow complex instructions without difficulty. We are confident that GL's mistakes in reading are not due to a failure to understand the task instructions. Additionally, GL has showed no evidence of a visual field neglect, nor was a visual field cut noted during his admission and rehabilitation at the hospital.

**Table 2**
Summary of reading errors during initial word-reading tasks.

| Word Class | Target | Response | Type of Error |
|---|---|---|---|
| Adjective | Glorious | Glorify | Morphological |
| Noun | Slavery | Slavement | Morphological |
| Noun | Movement | Moving | Morphological |
| Noun | Laughter | Laughing | Morphological |
| Noun | Insanity | Insaneness | Morphological |
| Noun | Hatred | Hating | Morphological |
| Noun | Bravery | Bravement | Morphological |
| Verb | Marry | Marriage | Morphological |
| Verb | Activation | Action | Morphological |
| Adjective | Weird | Odd | Semantic |
| Adjective | Soft | Soap[a] | Semantic |
| Adjective | Major | Corporal | Semantic |
| Noun | Kindness | Likeness | Semantic |
| Noun | Grief | Hurt | Semantic |
| Noun | Tolerance | Honorable | Semantic |
| Verb | Remember | Understand | Semantic |
| Noun | Rage | Cage | Visual |
| Adjective | Sparse | Spark | Visual |
| Noun | Infancy | Infanty | Visual |
| Noun | Elegance | Element | Visual |
| Noun | Defeat | Define | Visual |
| Noun | Advantage | Arranged | Visual |
| Verb | Detect | Detach | Visual |
| Verb | Illustrate | Illusion | Visual |
| Verb | Exist | Exit | Visual |
| Adjective | Rare | Hair | Visual/Phonological |

[a] The error of soap for soft likely comes from “Softsoap,” a North American brand name for soap products.

**Table 3**
Means of lexical-semantic variables for stimuli I.

| | Concrete | Abstract |
|---|---|---|
| Concreteness* | 569.56 | 327.78 |
| Letters | 6.36 | 6.38 |
| Phonemes | 5.18 | 5.38 |
| Syllables | 1.82 | 2.07 |
| Orthographic Frequency | 27.27 | 27.96 |
| Orthographic Neighbourhood Size | 1.47 | 1.07 |
| Familiarity | 515.82 | 512.09 |
| Age of Acquisition | 7.10 | 7.57 |
| Semantic Neighbourhood Density | .31 | .30 |
| Valence | 5.36 | 5.73 |
| Imageability* | 550.78 | 443.62 |
| Arousal* | 4.09 | 4.86 |

*Note.* Asterisks (*) indicate significant differences.

## 3. Materials

### 3.1. Stimuli I: stimulus development for words used in experiments 1 and 4

The stimulus set was developed using multiple databases to control for extraneous lexical variables. These databases include WINDSORS (Durda & Buchanan, 2008), WordMine2 (Durda & Buchanan, 2006), MRC Psycholinguistic Database (Coltheart, 1981), age of acquisition ratings (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012), and affective ratings (Warriner, Kuperman, & Brysbaert, 2013). The stimulus set consisted of 90 nouns (45 concrete and 45 abstract; see Appendix A). Concreteness was operationalized both subjectively and objectively through the use of the MRC Psycholinguistic Database (Coltheart, 1981). Concreteness values ranged from 100 to 700 ($M$ = 438). Concrete nouns were operationalized as stimuli with tangible referents that could be physically sensed, with values above the mean, whereas abstract nouns were operationalized as stimuli without tangible referents that could not be physically sensed, with values below the mean. The stimuli differed by concreteness, $t(71.93) = 26.77$, $p < .001$.

The remaining variables were used to tightly control any potential differences between concrete and abstract stimuli. These variables included word length, number of phonemes, number of syllables, orthographic frequency, orthographic neighbourhood size, familiarity, age of acquisition, semantic neighbourhood density (i.e., a measure of richness defined as the variability in the distribution of neighbouring words surrounding a target word's semantic neighbourhood; Durda & Buchanan, 2008), and valence. Imageability and arousal values were also obtained but not controlled. Consequently, concrete stimuli were easier to mentally visualize, $t(80.54) = 8.24$, $p < .001$, and abstract stimuli were more arousing, $t(88) = -3.32$, $p < .05$. Orthographic frequency values (Durda & Buchanan, 2006) were restricted to range between 3 and 67. Orthographic neighbourhood size was kept below 6. See Table 3 for a comparison between the various lexical-semantic variables for the concrete and abstract stimuli.

### 3.2. Stimuli II: stimulus development for words used in experiments 2, 3, and 4

The stimulus set was developed using WINDSORS (Durda & Buchanan, 2008) and Wordmine2 (Durda & Buchanan, 2006). The stimulus set consisted of 40 concrete word pairs (e.g., *moustache—beard*) and 40 abstract word pairs (e.g., *joy—sorrow*; see Appendix B). Concreteness was operationalized subjectively by the researchers consistent with the definition that concrete stimuli have direct sensory referents and can be easily visualized, while abstract stimuli do not have direct sensory referents and cannot be easily visualized (Schwanenflugel & Stowe, 1989). Concreteness ratings ($n$ = 50) collected confirmed this operationalization (see Appendix B for the ratings). Using a concreteness scale ranging from 1 to 7 and Altarriba, Bauer, and Benvenuto's (1999) instructions, half of the participants rated the individual word and half rated the word pair. Results showed that the mean concreteness rating of the abstract word pairs was 2.94 and the mean concreteness rating of the concrete word pairs was 6.32, and significantly different from one another, $F(1, 24) = 353.48$, $p < .001$. Moreover, while some ratings of the abstract *words* fell above the midpoint of 4, no abstract *word pair* was rated above the midpoint. The stimuli varied on semantic neighbourhood distance. Semantic neighbourhood distance is an ordinal measurement with the target word located X words away from its neighbour of interest. Close semantic neighbours were operationalized as having a distance of less than 50 words away from one another and distant semantic neighbours were operationalized having a distance of more than 200 words away from one another (e.g., *moustache* is the 2nd neighbour of *beard*, and *beard* is the 7th neighbour of moustache, making them close neighbours; *desk* is the 422$^{nd}$ neighbour of *carpet*, and *carpet* is the 361$^{st}$ neighbour of *desk*, making them distant neighbours). Half of the target word pairs were close semantic neighbours and half were distant semantic neighbours. The stimuli also varied on iconicity (i.e., whether word pairs were presented in the spatial relationships in which they typically occur, e.g., *moustache—beard*, or whether these relationships were reversed, e.g., *beard—moustache*). Half of the target word pairs were presented in an iconic relationship and half were presented in a reverse-iconic relationship. Iconicity was counterbalanced across Experiments 2 and 3 so that GL did not see the same word in iconic or reverse-iconic form in both tasks. The stimulus set for the semantic relatedness task also included 80 filler word pairs (see Appendix C) with no semantic relationship, as measured by WINDSORS, and no iconic relationship. Filler word pairs included both concrete and abstract words.

The target and filler word pairs were matched on word length. Target word pairs within the conditions were matched on average word length and average orthographic frequency. Orthographic frequency values (Durda & Buchanan, 2006) were restricted to range between 10 and 200. The concrete and abstract word pairs were matched on semantic neighbourhood distance. To avoid an alliteration effect, no two words in the pairs began with the same letter. Age of acquisition, imageability, emotional valence, and phonological neighbourhood density values were also obtained but not controlled. Age of acquisition was the older age associated with the word pair. As expected, the age of acquisition for concrete words pairs differed from the age of acquisition for abstract word pairs, $F(1, 79) = 14.05$, $p < .001$, such that abstract word pairs were acquired at a later age. Imageability ratings were found to differ, $F(1, 32) = 87.05$, $p < .001$, with mean imageability ratings for the concrete word pairs ($M = 151.02$) higher than mean imageability ratings for the abstract word pairs ($M = 108.69$). Emotional valence also differed, $F(1, 73) = 66.28$, $p < .001$, where abstract word pairs had greater differences[1] in emotional valence ($M = 3.17$) compared to the concrete word pairs ($M = 0.70$). Phonological neighbourhood density also differed, $t(39) = 3.04$, $p < .01$, where concrete word pairs had greater mean phonological neighbourhood densities ($M = 15.59$) than abstract word pairs ($M = 8.75$). See Table 4 for a comparison between the lexical variables for the concrete and abstract stimuli per condition.

## 4. Procedure

### 4.1. General experimental procedure overview

The study was approved by the University of Windsor Research Ethics Board. For each task, the procedure and expectations were described in detail to the participant and written informed consent was obtained. After the completion of each task, GL received $10 for participation. All tasks were conducted > 1 month apart. For Experiments 2 and 3 in which GL was presented with the same target word pairs, the presentation of the words was counterbalanced.

### 4.2. Computer-related tasks - procedure overview

The experiments were carried out in an individual testing room on a Dell PC computer with Windows XP operating system using DirectRT Software (Jarvis, 2012). The researcher remained in the room during testing to ensure comprehension was achieved throughout. For each task in Experiments 1, 2, and 3, GL was presented with stimuli and instructed to *not* read the words aloud and to make his response as quickly and accurately as possible. If he did not know the word, he was instructed to make a best guess. Concrete and abstract stimuli were presented individually in the center of the screen in a randomized order and remained on the screen until GL gave his response. GL was allowed a break halfway through each task. Reaction times and errors were recorded. The results of Experiment 1 were compared to the performance of controls (i.e., unimpaired readers) who completed the same task. The results of Experiments 2 and 3 were compared to the performance of controls who have completed the tasks from these experiments with identical stimuli (Malhi & Buchanan, 2018).

### 4.3. Experiment 1: concrete categorization task

Instructions were read aloud to the participant followed by a series of practice trials (with 4 concrete and 4 abstract stimuli) in which explicit feedback was provided. Concrete stimuli were described as words that could be physically sensed and that refer to a physical entity (e.g., *jacket, truck*). Abstract stimuli were characterized by words that could not be physically sensed and that do not refer to a physical entity (e.g., *democracy, crime*). GL was instructed to use his left index finger and to press the Z key if a word was concrete and the X key if a word was abstract. Between trials, he was told to place his left finger on the space bar. Stimuli were presented in all capital letters, with size 28 dark green coloured Times New Roman bold-faced font. The procedure was identical for controls except they were instructed to use both index fingers and to press the Z key if a word was concrete and the/key if a word was abstract.

### 4.4. Experiment 2: semantic relatedness task

Instructions were read aloud to the participant followed by a series of 4 practice trials (with 1 concrete and 1 abstract related word pair and 2 unrelated word pairs) in which explicit feedback was provided. GL was instructed to use his left index finger and to press the Z key if a word pair was related and the X key if a word pair was unrelated. Between trials, he was told to place his left finger on the space bar. See Appendix D for detailed instructions. Stimuli were presented in all capital letters, with size 24 turquoise

[1] Measuring the difference in emotional valence (operationalized as pleasantness; Warriner et al., 2013) between the words in this stimulus set was more informative than summing their valences. For example, the sum of emotional valence for *nose* and *tongue* was 11.79, and the sum of emotional valence for *beauty* and *ugly* was 10.05, indicating no difference in emotional valence between the concrete and abstract word pairs. However, the difference in emotional valence between *nose* and *tongue* was 0.79, and the difference between *beauty* and *ugly* was 5.11, suggesting a difference between the emotional valence differences of concrete and abstract word pairs. As the participant is attending to the relationship between the words, and these word pairs are presented vertically, the difference between one of the words being highly emotionally valenced (i.e., beauty) and the other being less emotionally valenced (i.e., ugly) indicates the pleasant-unpleasant relationships present within the abstract word pairs.

**Table 4**
Means of lexical variables per condition for stimuli II.

| | Concrete-Close | Concrete-Distant | Abstract-Close | Abstract-Distant |
|---|---|---|---|---|
| Letters | 10.9 | 10 | 12.15 | 11.9 |
| Orthographic Frequency | 37.81 | 35.14 | 44.81 | 41.73 |
| Age of Acquisition* | 6.15 | 6.19 | 7.68 | 8.12 |

*Note.* Asterisks (*) indicate significant differences.

coloured[2] Times New Roman bold-faced font. Word pairs were presented in a vertical position, with a distance of less than one inch between the middle of the two words.

### 4.5. Experiment 3: iconicity judgment task

Instructions were read aloud to the participant followed by a series of 4 practice trials (with 2 concrete and 2 abstract stimuli) in which explicit feedback was provided. GL was instructed to use his left index finger and to press the Z key if a word pair was in an iconic relationship and the X key if a word pair was in a reverse-iconic relationship. Between trials, he was told to place his left finger on the space bar. See Appendix E for detailed instructions. Stimuli were presented in all capital letters, with size 24 turquoise coloured Times New Roman bold-faced font. Word pairs were presented in a vertical position, with a distance of less than one inch between the middle of the two words.

### 4.6. Experiment 4: oral word-reading task

Words were presented individually on 11.6 cm by 11.1 cm sized paper cards. Words were printed in size 40 black coloured Arial font. The stimuli were those words used in the phonologically implicit tasks. The experiment proceeded in an individual room with two researchers. One researcher presented the stimuli directly in front of the participant in a random order and the other researcher recorded observations. GL was presented with words one at a time and instructed to read each word. If he hesitated or responded, “I don't know,” he was encouraged to make a best guess. Cards were not removed from view until he gave the correct response. The first phoneme was provided only after he had made multiple attempts at a word. All responses to stimuli were recorded verbatim.

## 5. Results

### 5.1. Data cleaning procedures

For the oral word-reading task, all items were retained in the analysis. GL's first attempt (i.e., response not including “I don't know”) was used to determine his accuracy rate and type of error made. For all computerized tasks, the same data cleaning procedure was applied: To remove outliers for each task, the mean reaction time (RT) was computed for correct items, and items that had RTs greater than 2 standard deviations above the mean were removed. An additional 9 items were removed for the RT analysis for the iconicity judgment task. For these 9 items, GL verbally asked for the definition for one of the words in each pair. Thus, reaction time was compromised and not accurately reflected by these items. There were no obvious patterns in the items removed with respect to the variables of interest. Error analysis was conducted when there was a > 10% error rate, and all items were retained in these analyses. All statistical tests for GL were conducted in IBM SPSS Statistics V22.0 Software. Uncontrolled lexical variables[3] were correlated with dependent variables to determine their influence on the outcome. For Experiments 2 and 3, age of acquisition and phonological neighbourhood density were uncontrolled for in the stimulus set. A Pearson correlation analysis was conducted on these variables and reaction time in these experiments. Age of acquisition was uncorrelated with reaction time in both Experiments 2 and 3, $r(67) < 0.01$, $p = .49$ and $r(55) = 0.02$, $p = .44$, respectively. Similarly, phonological neighbourhood density was uncorrelated with reaction time in both Experiments 2 and 3, $r(67) = 0.03$, $p = .78$ and $r(52) = 0.26$, $p = .05$, respectively. Word length and age of acquisition were also uncontrolled in Experiment 4. A point biserial correlation was used to determine the relation between age of acquisition and errors, and a weak significant correlation was obtained with more errors for later acquired words $r(248) = 0.43$, $p < .001$. An identical analysis was conducted to determine the relation between word length and errors, and a significant, albeit very small magnitude, correlation was obtained with more errors for longer words, $r(248) = 1.46$, $p < .05$. The data were split by concreteness to determine whether the uncontrolled lexical variables influenced the outcome similarly across each level of the independent variable. This was found to be the case with weak correlations found between age of acquisition and errors for concrete ($r = 0.33$) and abstract ($r = 0.43$) words, ($ps < .001$), and weak, non-significant correlations between word length and errors for concrete and abstract words ($ps > .05$). Thus, these variables were not included in further analyses. Data are available through

[2] There was no scientific premise for the choice of font colour and the font colour was not analyzed or expected to affect the results.

[3] Although imageability, arousal, and valence were uncontrolled for in the stimuli, imageability is thought to correlate with the concreteness of a word, and emotional salience (i.e., arousal and valence) is thought to correlate with the abstractness of a word. Thus, it is expected that these variables would influence reaction time.

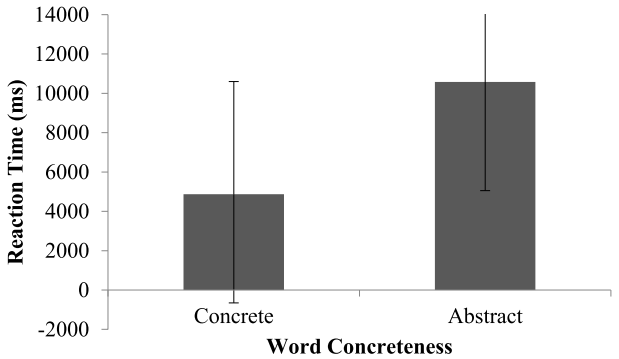


**Fig. 4.** Reaction times for the concrete categorization task. GL was faster to correctly categorize concrete words than abstract words.

Mendeley Data.

Control data in the concrete categorization task was analyzed using R (R Core Team, 2016) version 3.4.3 and the lme4 and lmerTest packages (Kuznetsova, Brockhoff, & Christensen, 2013). There were no invalid responses (< 300 ms) and no participants with accuracy rates < 70%. Eight words had accuracy rates < 70% and were removed from the analysis (128 observations). Errors were removed (74 observations; 6% of the remaining data).

### 5.2. Experiment 1: concrete categorization task

#### 5.2.1. RT analysis

A one-way between-subjects ANOVA was conducted to compare the effect of concreteness on reaction times for categorizing concrete and abstract stimuli. According to Levene's test, the assumption of homogeneity of variance was not violated, $F(1, 63) = 0.39$, $p = .54$. There was an effect of concreteness on reaction time, $F(1, 63) = 15.93$, $p < .001$, in which GL was faster to correctly categorize concrete ($M = 4871.15$, $SD = 5729.73$) than abstract ($M = 10580.04$, $SD = 5527.44$) stimuli (see Fig. 4). Correct responses from the control data were analyzed in a linear mixed effects analysis. RTs were log transformed. As fixed effects, the factor concreteness was considered for the model. As random effects, subjects with random slopes for concreteness and items with random slopes were considered for the model. Using forward and backward fitting, variables were removed, and the final model included concreteness as a fixed effect, random slopes for concreteness by subject, and random slopes for item. P-values were obtained for the fixed effects using the lmerTest package with Satterthwaite approximations to degrees of freedom (Kuznetsova et al., 2013). After the model was fitted, data were trimmed and outliers with a standardized residual at a distance greater than 2.5 SD from 0 were excluded. This resulted in the removal of 37 observations (2.94% of the data). There was a main effect of concreteness, $b = -0.08$, $t(28.27) = -2.19$, $p < .05$, in which controls were faster to correctly categorize concrete ($M = 946.51$, $SD = 382.50$) than abstract ($M = 1026.19$, $SD = 452.60$) stimuli.

#### 5.2.2. Error analysis

A Chi-square test of independence was calculated comparing the frequency of errors among concrete and abstract stimuli. An interaction was found, $\chi(1) = 9.36$, $p < .05$, $\phi = 0.32$. GL was more likely to categorize abstract stimuli as concrete (21%) than concrete stimuli as abstract (7%), consistent with the RT analysis. See Table 5 for an error count for the concrete and abstract stimuli. For control data accuracy, the binomial dependent variable (i.e., correct or incorrect) was analyzed using a mixed logit model (generalized linear mixed model; Jaeger, 2008). As fixed effects, the factor concreteness was considered for the model. The main effect of concreteness approached significance, $b = -0.69$, $z = -1.93$, $p = .05$, in terms of categorizing abstract stimuli as concrete (7%) and concrete stimuli as abstract (4%).

### 5.3. Experiment 2: semantic relatedness task

#### 5.3.1. RT analysis

A 2 (concreteness: concrete and abstract) X 2 (semantic distance: close and distant) X 2 (iconicity: iconic and reverse-iconic) factorial ANOVA was conducted with RT as the dependent variable. The assumption of homogeneity of variance was violated according to Levene's Test, $F(7, 61) = 4.33$, $p < .05$. However, ANOVA is robust against moderate violations, particularly when group

**Table 5**
Error count per category for concrete categorization task.

| Concreteness | False | True |
|---|---|---|
| Concrete | 6 | 39 |
| Abstract | 19 | 26 |

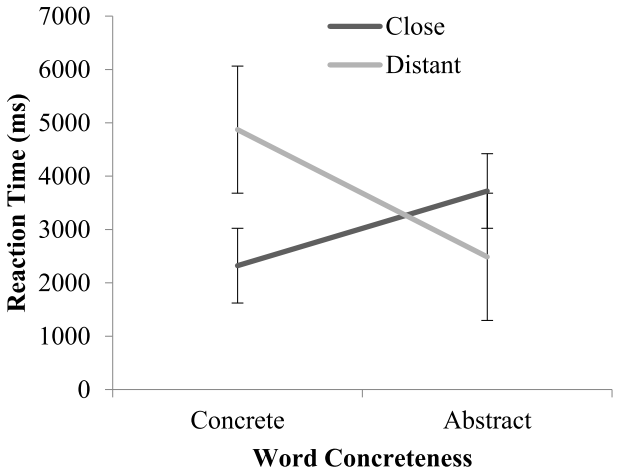


**Fig. 5.** Reaction times for reverse-iconic word pairs in the semantic relatedness task. GL was slower to correctly identify that concrete, distant word pairs were related in their reverse-iconic form.

sizes are relatively equal, which was the case (between $n = 33$ and $n = 36$). There were no main effects or two-way interactions ($p$'s $> 0.05$). There was a three-way interaction, $F(1, 61) = 4.49$, $p < .05$. To disentangle the three-way interaction, the data were split by iconicity. For reverse-iconic word pairs, there was an interaction between concreteness and semantic distance, $F(1, 32) = 6.86$, $p < .05$. For concrete, reverse-iconic stimuli, visual inspection of the means and graph showed that GL was slower at making a correct decision for distant word pairs ($M = 4872.40$, $SD = 2959.32$) relative to close word pairs ($M = 2321.38$, $SD = 564.94$). In contrast, abstract, reverse-iconic stimuli exhibited the reverse effect, albeit less dramatically, with GL demonstrating a relatively slower response for close word pairs ($M = 3722.11$, $SD = 2833.36$) than distant word pairs ($M = 2488.11$, $SD = 727.64$). See Fig. 5 for a visual depiction of the interaction. For iconic word pairs, there were no significant results to report ($p$'s $> 0.05$).

### 5.4. Experiment 3: iconicity judgment task

#### 5.4.1. RT analysis

A 2 (concreteness: concrete and abstract) X 2 (semantic distance: close and distant) X 2 (iconicity: iconic and reverse-iconic) factorial ANOVA was conducted with RT as the dependent variable. A main effect of concreteness was found, $F(7, 49) = 10.75$, $p < .05$, in which GL was faster to respond to abstract word pairs ($M = 4425.10$, $SD = 2182.92$) than concrete word pairs ($M = 6815.04$, $SD = 2996.95$). A main effect of iconicity was also found, $F(7, 49) = 4.14$, $p < .05$, in which GL was faster at making decisions for iconic word pairs ($M = 4844.39$, $SD = 2095.83$) relative to reverse-iconic word pairs ($M = 6407.04$, $SD = 3385.73$). See Table 6 for a summary of the mean differences. The main effect of distance, two-way interactions, and three-way interaction were all non-significant ($p$'s $> 0.05$).

### 5.5. Experiment 4: oral word-reading task

For the analyses, the two stimulus sets used in the phonologically implicit tasks were combined. The concrete and abstract stimuli in the combined stimulus set were matched for orthographic frequency ($p > .05$), although they differed by word length, $t(248) = -2.47$, $p < .05$ and by age of acquisition, $t(248) = -5.69$, $p < .001$. In both instances, abstract stimuli had longer word lengths and a higher age of acquisition relative to concrete stimuli. See Table 7 for a summary of the mean lexical variables for concrete and

**Table 6**
Summary of mean RT (ms) Differences for iconicity judgment task.

| Condition | Mean | Standard Deviation |
|---|---|---|
| *Concreteness* | | |
| Concrete | 6815.04 | 2996.95 |
| Abstract | 4425.10 | 2182.92 |
| *Distance* | | |
| Close | 5564.55 | 2834.92 |
| Distant | 5549.54 | 2904.07 |
| *Iconicity* | | |
| Iconic | 4844.39 | 2095.83 |
| Reverse-Iconic | 6407.04 | 3385.73 |

**Table 7**
Means of lexical variables for combined concrete and abstract stimuli.

| | Concrete | Abstract |
|---|---|---|
| Letters* | 5.64 | 6.14 |
| Orthographic Frequency | 32.27 | 37.54 |
| Age of Acquisition* | 6.02 | 7.34 |

*Note.* Asterisks (*) indicate significant differences.

abstract stimuli. A Chi-square test of independence was calculated comparing the frequency of reading errors among concrete and abstract stimuli for the combined stimulus set. An interaction was found, $\chi(1) = 9.78$, $p < .01$, $\phi = 0.20$, such that GL made more errors when reading abstract words (18%) than concrete words (9%). See Table 8 for examples of reading errors.

## 6. Discussion

The objective of this study was to test the application of FIT (Buchanan et al., 2003) to concrete and abstract stimuli using implicit and explicit task demands. FIT proposes that in deep dyslexia, errors to abstract words are the result of an impairment in phonological output lexicon selection rather than a semantic deficit for abstract words. Further, FIT posits a dissociation between explicit production and implicit access of representations, such that production can be compromised whilst access of representations is intact. Therefore, it was hypothesized (see Table 1 for a summary of hypotheses) that in phonologically implicit tasks where controls demonstrate either concreteness or abstractness effects, a participant with deep dyslexia would similarly show concreteness or abstractness effects. However, for explicit tasks where production is involved, a participant with deep dyslexia would only show concreteness effects due to difficulty with abstract word production, indicative of their difficulty with phonological output lexicon selection for abstract words. It is important to note that the performance of the participant with deep dyslexia was not expected to be the same as controls, as GL was expected to be slower and make more errors overall. However, the same pattern of performance between GL and controls was expected, whether for concreteness effects (Friederici et al., 2000) or abstractness effects (Malhi & Buchanan, 2018). The results of the experiments generally supported these predictions.

Experiment 1 tapped into implicit access of concrete versus abstract representations. As was observed in controls and as previously shown in controls (Friederici et al., 2000), GL demonstrated concreteness effects, such that he was faster in categorizing concrete words compared to abstract words and he had fewer errors when categorizing concrete words. While the implicit task in Experiment 1 replicated the concreteness effects that are observed in controls, the implicit tasks in Experiments 2 and 3 sought to replicate the abstractness effects that are observed in controls who have completed these tasks with identical stimuli (Malhi & Buchanan, 2018). In conjunction with the results of Experiment 1, if abstractness effects could also be replicated with GL in Experiments 2 and 3, this would support that implicit representations are intact in deep dyslexic readers, consistent with FIT. In addition, it would discredit the view that deep dyslexia involves a semantic deficit for abstract words. Indeed, abstractness effects were observed in Experiment 3. Experiment 2 tapped into implicit access of the semantic relatedness of concrete and abstract word pairs. When controls complete the semantic relatedness task from Experiment 2, in addition to an advantage for abstract word pairs, a finding is that word pairs that are close semantic neighbours are determined to be related faster than word pairs that are distant semantic neighbours (Malhi & Buchanan, 2018). GL showed this pattern of results for concrete word pairs presented in reverse-iconic form. However, he showed an opposite pattern of results, albeit less dramatically, for abstract reverse-iconic word pairs. Thus, semantic distance provided an advantage for concrete words, but abstract words did not benefit from being paired with a close semantic neighbour. Moreover, Malhi and Buchanan (2018) noted that opposite, but related, word pairs tended to be more prevalent with abstract stimuli (e.g., *joy—sorrow*) whereas concrete stimuli were related but not opposites (e.g., *shirt—pants*). Given that abstract word pairs were related by being opposites, this may have limited the role of semantic distance in facilitating processing for GL. Experiment 3 tapped into implicit access of the spatial representations of concrete and abstract word pairs. When controls complete the iconicity judgment task from Experiment 3, in addition to an advantage for abstract word pairs, a finding is that word pairs that are presented in an iconic form are responded to faster than word pairs presented in a reverse-iconic form (Malhi & Buchanan, 2018). GL replicated both findings; he responded faster to abstract word pairs compared to concrete word pairs and iconic word pairs compared to reverse-iconic word pairs. The replication of the abstractness effect in Experiment 3 provides support in favour of FIT and provides support against deep dyslexic readers having a semantic deficit for abstract words. The iconicity judgment task from Experiment 3 is a deep processing task that subsumes access of the semantic representations of abstract words, as the semantic representations must be accessed first in order to then mentally manipulate their spatial locations. While we had hypothesized that GL would demonstrate abstractness effects in both Experiments 2 and 3 as do controls, GL only showed an abstractness effect for Experiment 3. However, controls have a larger abstractness effect for the iconicity judgment task than they do for the semantic relatedness task, so GL replicated the larger effect. In the introduction, we discussed various theories that explain an impairment for abstract word processing in deep dyslexia. The results of Experiments 1–3 are inconsistent with theories that suggest a semantic deficit for abstract words, such as Coltheart's (1980b) right hemisphere hypothesis, Plaut and Shallice's (1993) connectionist model, and Morton and Patterson's (1980) dual-route model. However, the results are consistent with theories that suggest intact semantic processing of concrete and abstract words in deep dyslexia including FIT (Buchanan et al., 2003) and Newton and Barry's NICE model (1997).

**Table 8**
Summary of reading errors during oral word-reading task.

| Stimulus Set | Target | Response | Type of Error |
|---|---|---|---|
| Stimuli I | Product | Production | Morphological |
| | Warmth | Warmness | Morphological |
| | Warehouse | House | Morphological |
| | Thrill | Drill | Phonological |
| | Branch | Mranch | Phonological |
| | Sensation | Censorship | Phonological |
| | Cliff | Fall | Semantic |
| | Portrait | Frame | Semantic |
| | Dagger | Danger | Semantic |
| | Error | Erase | Semantic |
| | Luxury | Deluxe | Semantic |
| | Garment | Garnit | Visual |
| | Ivory | Ivy | Visual |
| | Carriage | Car | Visual |
| | Powder | Power | Visual |
| | Corridor | Courtmenship | Visual |
| | Stool | Tool | Visual |
| | Ambition | Attitude | Visual |
| | Approval | Arrival | Visual |
| | Attitude | Longitude | Visual |
| | Burden | Burial | Visual |
| | Chaos | Chase | Visual |
| | Envy | Ivy | Visual |
| | Equality | Equit | Visual |
| | Frenzy | Freezing | Visual |
| | Myth | Match | Visual |
| | Paradise | Parade | Visual |
| | Estate | State | Visual/Phonological |
| | Skill | Skull | Visual/Phonological |
| | Slumber | Bugging | Other |
| | Removal | Moatley | Other |
| Stimuli II | Runway | Running | Morphological |
| | Therapist | Therapy | Morphological |
| | Crooked | Crooking | Morphological |
| | Hiker | Hiking | Morphological |
| | Handle | Hand | Morphological |
| | Boredom | Boarding | Morphological |
| | Ally | Aily | Phonological |
| | Gain | Jane | Phonological |
| | Heel | Knee | Semantic |
| | Enemy | Lost | Semantic |
| | Foam | Soap | Semantic |
| | Jeans | Pants | Semantic |
| | Defeat | Defy | Visual |
| | Dim | Dime | Visual |
| | Excitement | Exercise | Visual |
| | Reject | Eject | Visual |
| | Smooth | Mouth | Visual |
| | Tenant | Tendant | Visual |
| | Borrow | Broken | Visual |
| | Disagree | Disorganize | Visual |
| | Host | Lost | Visual |
| | Increase | Inside | Visual |
| | Tainted | Painting | Visual/Morphological |
| | Scarce | Scared | Visual/Morphological |
| | Meek | Peak | Visual/Phonological |
| | Bless | Bliss | Visual/Semantic |
| | Tray | Fast | Other |

While FIT proposes that access of semantic representations is intact in deep dyslexia, FIT also proposes that explicit production is compromised in deep dyslexia. Therefore, the second component to our test of FIT with concrete and abstract words involved incorporating a task with explicit production demands. In Experiment 4, GL read aloud concrete and abstract words from two different stimulus sets, although the stimuli used were identical to the stimuli used in the implicit tasks. Results showed a concreteness effect whereby GL made more errors when reading aloud abstract words than concrete words. Notably, GL showed this disadvantage for reading aloud abstract words, despite having an advantage for responding to abstract words when explicit production was not required, i.e., in Experiment 3. Therefore, consistent with FIT, both concreteness and abstractness effects were

observed, with the type of effect observed depending on whether the task required explicit production or implicit access and mirroring performance patterns found in controls on phonologically implicit tasks. A confound in Experiment 4 is that abstract stimuli had longer word lengths and a higher age of acquisition relative to concrete stimuli. While these extraneous lexical variables may have influenced explicit production in addition to concreteness, the critical test of FIT for this study was to demonstrate that deep dyslexia does not reflect a semantic deficit for abstract words. Indeed, GL demonstrated abstractness effects with respect to comprehension for these same words (i.e., longer abstract words and words with higher age of acquisition ratings) in Experiment 3. Nevertheless, future research may be useful in understanding this interaction between extraneous lexical variables and concreteness effects in reading aloud in deep dyslexia. For example, Crutch and Warrington (2005) report that concrete words are related by similarity and abstract words are related by association. As such, another confound in our study may result from the concrete and abstract word pairs being related by both similarity and association. However, while this may be a confound, it is not incompatible with our notion of abstract words having fewer constraints in the phonological output lexicon. That is, abstract words may have more associates (i.e., semantic neighbours) because of both similarity and association.

Overall, the results of these experiments are consistent with FIT. GL did not appear to have a semantic deficit for abstract words. In addition, GL had intact implicit access to concrete and abstract words, as highlighted by his similar performance to controls on phonologically implicit tasks. However, when explicit production was required, GL's performance diverged from that of controls. The concreteness effect observed in the explicit production task could be attributed to an impairment in phonological output lexicon selection for abstract words. Determining the locus of impairment in deep dyslexia is a substantial area for further research. Such research contributes to the basis of informing treatment in persons with deep dyslexia. For example, Ablinger and Radach (2016) found that a new eye movement supported intervention combining lexical and segmental strategies for deep dyslexia especially improved phonological output lexicon abilities for concrete words.

In addition, some researchers (see Riley & Thompson, 2010) have suggested a modified version of FIT, with failure of inhibition beginning at the level of semantic processing. In their study, Riley and Thompson (2010) did not find a semantic typicality effect when they asked deep dyslexic readers to indicate whether a target word belonged to a semantic category. Given their data, a modification of FIT to account for an impairment in semantic processing may seem reasonable. However, these researchers did not test semantic processing without phonological output lexicon demands. Their semantic categorization task required the deep dyslexic reader to hold words in their phonological working memory before making a response. An alternative explanation for their data, and one that can enrich FIT further, is the link between phonological working memory deficits and problems with explicit phonological output lexicon selection in deep dyslexia. The investigation of this link may be a fruitful area for future research on deep dyslexia.

## Acknowledgments

We thank GL for his ongoing participation in our research. This work was supported by Social Sciences and Humanities Research Council scholarships awarded to SKM (767-2016-1440) and TM (752-2018-2102) and a Words in the World Partnership Grant awarded to LB (895-2016-1008). There is no perceived conflict of interest. The work described here has not been published previously, is not under consideration for publication elsewhere, if accepted it will not be published elsewhere in the same form, and its publication has been approved by all authors.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jneuroling.2018.11.001.

## References

Ablinger, I., & Radach, R. (2016). Diverging receptive and expressive word processing mechanisms in a deep dyslexic reader. *Neuropsychologia, 81*, 12–21. https://doi.org/10.1016/j.neuropsychologia.2015.11.023.

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers, 31*(4), 578–602. https://doi.org/10.3758/BF03200738.

Barry, C., & Gerhand, S. (2003). Both concreteness and age-of-acquisition affect reading accuracy but only concreteness affects comprehension in a deep dyslexic patient. *Brain and Language, 84*, 84–104. https://doi.org/10.1016/S0093-934X(02)00522-9.

Bose, A., Colangelo, A., & Buchanan, L. (2011). Effect of phonetic complexity on word reading and repetition in deep dyslexia. *Journal of Neurolinguistics, 25*, 435–444. https://doi.org/10.1016/j.jneuroling.2011.01.004.

Boumaraf, A., & Macoir, J. (2016). The influence of visual word form in reading: Single case study of an Arabic patient with deep dyslexia. *Reading and Writing, 29*, 137–158. https://doi.org/10.1007/s11145-015-9583-y.

Buchanan, L., Hildebrandt, N., & MacKinnon, G. E. (1994). Implicit phonological processing in deep dyslexia: A rowse is implicitly a rose. *Journal of Neurolinguistics, 8*, 163–182. https://doi.org/10.1016/0911-6044(96)00001-2.

Buchanan, L., Hildebrandt, N., & MacKinnon, G. E. (1996). Phonological processing of nonwords in deep dyslexia: Typical and independent. *Journal of Neurolinguistics, 9*, 113–133. https://doi.org/10.1016/0911-6044(96)00001-2.

Buchanan, L., Hildebrandt, N., & MacKinnon, G. E. (1999). Phonological processing in acquired deep dyslexia re-examined. In R. Klein, & P. McMullen (Eds.). *Converging methods for understanding reading and dyslexia.* Cambridge, MA, US: The MIT Press.

Buchanan, L., Kiss, I., & Burgess, C. (2000). Phonological and semantic information in word and nonword reading in a deep dyslexic patient. *Brain and Cognition, 43*, 65–68. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10857665.

Buchanan, L., McEwen, S., Westbury, C., & Libben, G. (2003). Semantics and semantic errors: Implicit access to semantic information from words and nonwords in deep dyslexia. *Brain and Language, 84*(1), 65–83. https://doi.org/10.1016/S0093-934X(02)00521-7.

Colangelo, A., & Buchanan, L. (2005). Semantic ambiguity and the failure of inhibition hypothesis as an explanation of reading errors in deep dyslexia. *Brain and Cognition, 57*, 39–42. https://doi.org/10.1016/j.bandc.2004.08.018.

Colangelo, A., & Buchanan, L. (2006). Implicit and explicit processing in deep dyslexia: Semantic blocking as a test for failure of inhibition in the phonological output

lexicon. *Brain and Language, 99*(3), 258–271. https://doi.org/10.1016/j.bandl.2005.07.048.
Colangelo, A., & Buchanan, L. (2007). Localizing damage in the functional architecture: The distinction between implicit and explicit processing in deep dyslexia. *Journal of Neurolinguistics, 20*, 111–144. https://doi.org/10.1016/j.jneuroling.2006.08.001.
Colangelo, A., Stephenson, K., Westbury, C., & Buchanan, L. (2003). Word associations in deep dyslexia. *Brain and Cognition, 53*(2), 166–170. https://doi.org/10.1016/S0278-2626(03)00102.
Colangelo, A., Westbury, C., & Buchanan, L. (2004). Deep dyslexia and semantic errors: A test of the failure of inhibition hypothesis using a semantic blocking paradigm. *Brain and Cognition, 54*, 232–234. https://doi.org/10.1016/j.bandc.2004.02.016.
Coltheart, M. (1980a). Deep dyslexia: A review of the syndrome. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.). *Deep dyslexia*. London: Routledge & Kegan Paul.
Coltheart, M. (1980b). Deep dyslexia: A right-hemisphere hypothesis. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.). *Deep dyslexia*. London: Routledge & Kegan Paul.
Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33*(4), 497–505. https://doi.org/10.1080/14640748108400805.
Coltheart, M., Patterson, K. E., & Marshall, J. C. (1980). *Deep dyslexia.* London: Routledge & Kegan Paul.
Crutch, S. J. (2006). Qualitatively different semantic representations for abstract and concrete words: Further evidence from the semantic reading errors of deep dyslexic patients. *Neurocase, 12*, 91–97. https://doi.org/10.1080/13554790500507172.
Crutch, S. J., & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain, 128*, 615–627. https://doi.org/10.1093/brain/awh349.
Durda, K., & Buchanan, L. (2006). *WordMine2.* Available from: http://web2.uwindsor.ca/wordmine.
Durda, K., & Buchanan, L. (2008). WINDSORS: Windsor improved norms of distance and similarity of representations of semantics. *Behavior Research Methods, 40*(3), 705–712. https://doi.org/10.3758/BRM.40.3.705.
Friederici, A. D., Opitz, B., & von Cramon, D. Y. (2000). Segregating semantic and syntactic aspects of processing in the human brain: An fMRI investigation of different word types. *Cerebral Cortex, 10*(7), 698–705. https://doi.org/10.1093/cercor/10.7.698.
Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446. https://doi.org/10.1016/j.jml.2007.11.007.
Jarvis, B. G. (2012). *DirectRT (version 2012) [computer software].* New York, NY: Empirisoft Corporation.
Katz, R. B., & Lanzoni, S. M. (1997). Activation of the phonological lexicon for reading and object naming in deep dyslexia. *Brain and Language, 58*, 46–60. https://doi.org/10.1006/brln.1997.1766.
Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). *R Package Version, 2*(6).
Malhi, S. K., & Buchanan, L. (2018). A test of the symbol interdependency hypothesis with both concrete and abstract stimuli. *PLoS One, 13*(3), e0192719. https://doi.org/10.1371/journal.pone.0192719.
Marelli, M., Aggujaro, S., Molteni, F., & Luzzatti, C. (2012). The multiple-lemma representation of Italian compound nouns: A single case study of deep dyslexia. *Neuropsychologia, 50*(5), 852–861. https://doi.org/10.1016/j.neuropsychologia.2012.01.021.
Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research, 2*, 1975–1999. https://doi.org/10.1007/BF01067101.
McClelland, J. L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.). *Attention and performance XII: The psychology of reading* (pp. 3–36). Hillsdale, NJ: Erlbaum.
Morton, J., & Patterson, K. E. (1980). A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.). *Deep dyslexia*. London: Routledge & Kegan Paul.
Newton, P. K., & Barry, C. (1997). Concreteness effects in word production but not word comprehension in deep dyslexia. *Cognitive Neuropsychology, 14*(4), 481–509. https://doi.org/10.1080/026432997381457.
Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*, 255–287. https://doi.org/10.1037/h0084295.
Plaut, D., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology, 10*, 377–500. https://doi.org/10.1080/02643299308253469.
R Core Team (2016). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.
Riley, E. A., & Thompson, C. K. (2010). Semantic typicality effects in acquired dyslexia: Evidence for semantic impairment in deep dyslexia. *Aphasiology, 24*(6–8), 802–813. https://doi.org/10.1080/02687030903422486.
Ripamonti, E., Aggujaro, S., Molteni, F., Zonca, G., Frustaci, M., & Luzzatti, C. (2014). The anatomical foundations of acquired reading disorders: A neuropsychological verification of the dual-route model of reading. *Brain and Language, 134*, 44–67. https://doi.org/10.1016/j.bandl.2014.04.001.
Sandberg, C., & Kiran, S. (2014). Analysis of abstract and concrete word processing in persons with aphasia and age-matched neurologically healthy adults using fMRI. *Neurocase, 20*(4), 361–388. https://doi.org/10.1080/13554794.2013.770881.
Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language, 27*, 499–520. https://doi.org/10.1016/0749- 596X(88)90022-8.
Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 82–102. https://doi.org/10.1037/0278-7393.9.1.82.
Schwanenflugel, P. J., & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly, 24*(1), 114–126. https://doi.org/10.2307/748013.
Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x.
Warrington, E. K., & Shallice, T. (1979). Semantic access dyslexia. *Brain, 102*, 43–63. https://doi.org/10.1093/brain/102.1.43.