DISCOURSE PROCESSES
http://dx.doi.org/10.1080/0163853X.2017.1289794

Routledge
Taylor & Francis Group

Can Differences in Word Frequency Explain Why Narrative Fiction Is a Better Predictor of Verbal Ability than Nonfiction?

Graham A. McCreath[a], Cormac M. J. Linehan[b], and Raymond A. Mar[a]

[a]Department of Psychology York University, Toronto, Ontario, Canada; [b]Department of Linguistics York University, Toronto, Ontario, Canada

ABSTRACT

Individuals who read more tend to have stronger verbal skills than those who read less. Interestingly, what you read may make a difference. Past studies have found that reading narrative fiction, but not expository nonfiction, predicts verbal ability. Why this difference exists is not known. Here we investigate one possibility: whether fiction texts contain more of the words typically evaluated by verbal ability measures compared to nonfiction texts. We employed corpus linguistic analyses to compare the frequency with which commonly tested SAT words appeared in both fiction and nonfiction texts, for 3 different corpora. Differences in SAT word frequency between the two genres were found to be negligible across all corpora. As a result, we conclude that there is little evidence that differences in word content between fiction and nonfiction texts can account for their differential relation to verbal ability. Other possible explanations are proposed for future research.

Introduction

The observation that individuals who read more have better verbal abilities is among one of the most robust findings in reading research. For example, a meta-analysis of 99 studies revealed moderate to strong correlations between reading, measured as lifetime print exposure (Martin-Chang & Gould, 2008; Stanovich & West, 1989), and several different indices of linguistic competence (Mol & Bus, 2011). (As with all correlational data, causal inferences cannot be supported and alternative third-variable explanations may exist.) These associations were present across a wide range of periods in the lifespan, from preschoolers and kindergartners, through children in grades 1 to 12, up to and including college- and university-aged students (N = 7,669; Mol & Bus, 2011). Notably, print exposure remains a strong predictor of linguistic ability even after controlling for related factors such as age, education, reading comprehension, and intelligence for both children (Allen, Cipielewski, & Stanovich, 1992; Cipielewski & Stanovich, 1992; Cunningham & Stanovich, 1991; Davidse, de Jong, Bus, Huijbregts, & Swaab, 2011; Ecalle & Magnan, 2008) and adults (Burt & Fury, 2000; Cunningham & Stanovich, 1997; Sparks, Patton, & Murdoch, 2014; Stanovich & Cunningham, 1992; West, Stanovich, & Mitchell, 1993).[1]

Given the robust association between reading and language abilities, it is important to consider more nuanced questions, such as whether unique genres of text might differentially predict verbal ability. To this end, there is growing evidence that literary skills are more closely associated with exposure to narrative fiction compared with expository nonfiction. For example, in a study of sixth grade children, self-reported reading of fiction books was found to predict word identification and

CONTACT Raymond A. Mar mar@yorku.ca Department of Psychology, York University, 4700 Keele St., Toronto, ON, M3J 1P3, Canada.

[1]Note that by predict, here and throughout, we mean that one can make an educated guess about some variable based on knowledge of another, as is the case with correlation and regression. We are making no causal claims in our use of this word.

reading comprehension, but no such association was observed for nonfiction books, magazines, newspapers, or comics (Spear-Swerling, Brucker, & Alfano, 2010). In a separate study on 11- to 16-year-old students, time spent reading fiction predicted a variety of reading-based skills (e.g., word attack, reading comprehension), unlike time spent reading factual books, school textbooks, and a variety of shorter texts (e.g., magazines, e-mails, social media content; McGeown, Duncan, Griffiths, & Stothard, 2015). This advantage for fiction in predicting verbal ability has also been demonstrated in adults. Mar and Rain (2015) conducted a series of three studies in which they assessed verbal ability using three different subtests of the verbal section of the SAT (formerly the Scholastic Aptitude Test): analogies, sentence completion, and reading comprehension. They found that exposure to fiction—as assessed either by self-reported reading habits or by a recognition-based print-exposure measure—was a consistent predictor of verbal ability whereas exposure to nonfiction was not. Moreover, fiction demonstrated an association even after controlling for demographic variables and exposure to nonfiction texts. A fourth study on synonyms found similar results, demonstrating that this advantage for fiction over nonfiction is robust and replicable (total *N* across all 4 studies = 960). In light of the fact that these studies examined diverse age groups with relatively large samples and used various broad and indirect measures of reading behavior, this preferential prediction of verbal abilities by fiction over nonfiction appears to be generalizable across both populations and texts (i.e., not tied to particular types of participants or the specific texts that they read).

But why might exposure to narrative fiction predict verbal ability better than expository nonfiction? Previous studies have identified differences in the content of fiction and nonfiction texts (Gardner, 2004; Nation & Waring, 1997). Perhaps these content differences can explain why the two genres are differentially associated with verbal ability. For example, fiction texts might be more likely to contain the kind of words that are often encountered in standardized tests of verbal ability, relative to nonfiction texts. To evaluate this possibility, we examined large corpora that included both fiction and nonfiction texts, comparing the frequency with which commonly tested words appear in each genre. To map this examination as closely as possible to the findings of Mar and Rain (2015), we chose to use the SAT as a word base for words typically assessed in standardized tests of verbal ability. We then performed a corpus linguistic analysis to compare the frequency with which these words appeared in both fiction and nonfiction texts.

Methods

Common SAT words

For the purpose of this study, we selected a list of words published by the test preparation company Ivy Global, the "500 Most Important Words for the SAT." This list was chosen as a sample of commonly tested words as it freely available (as opposed to proprietary) and contains a comprehensive survey of tested words. The list provides multiple derivational suffixes (e.g., satire, satirical) for many words and we included each form in our own list, resulting in a total of 598 SAT words.

Text corpora

We analyzed the frequency with which SAT words appeared in fiction and nonfiction genres using three large corpora of texts: the Corpus of Contemporary American English (COCA), the British National Corpus (BNC), and the Strathy Corpus of Canadian English (hereafter Strathy). We selected these corpora because they include detailed information regarding the genre of the texts, which is absent from many other sources (e.g., Google Books). Moreover, by using these three corpora we are able to examine whether any results observed replicate across the three different corpora, which are each based on a different geographically linked form of English. As a result, we are able to examine both the robustness and generalizability of any results.

COCA. The COCA contains 534 million words, 80% of which are derived from written sources published between 1990 and 2015 (Davies, 2008). The remainder of the corpus is made up of transcripts of spoken dialogue. The written portion of COCA includes a category for the Fiction genre and three relevant Nonfiction genres for comparison: Academic journals, Popular magazines, and Newspapers. Each genre contains approximately 105 million words derived from between 24,000 and 67,000 different texts. The Fiction category includes short stories, plays, book chapters, and movie scripts. The categories of Academic journals and Popular magazines consist of articles drawn from nearly 100 different peer-reviewed academic journals and general interest magazines spanning a wide range of domains. The Newspapers category is comprised of various sections of 10 large American newspapers.

BNC. The BNC contains 96 million words, 90% of which are from texts published from the latter part of the 20th century with the remaining 10% comprised of spoken dialogue transcripts (Davies, 2004). Like COCA, the BNC contains a category for Fiction only, which consists of novels and poetry, as well as the Nonfiction categories of Academic journals, Popular magazines, and Newspapers. The BNC also includes another Nonfiction category not found within COCA, Non-academic texts, which consists mostly of nonfiction books. The number of words in each genre ranges from 7 to 16 million words and the number of texts for each genre ranges from 211 to 530.

Strathy. The Strathy Corpus contains 50 million words from Canadian texts published between 1921 and 2011, with 10% of the content based on spoken language (Strathy Language Unit, 2013). Strathy includes the same genres as COCA as well as an additional Nonfiction category designated as "Nonfiction texts," which consists mostly of essays and nonfiction books. The number of words in each genre ranges from 2 to 15 million words and the number of texts for each genre from 105 to 2,837.

Procedure

We queried each corpus for the frequency of each of the 598 SAT words within each genre using a web interface created and maintained by Mark Davies at Brigham Young University (http://corpus.byu.edu/). Frequencies were set as frequency per million words to control for the different sizes of the three corpora and the size of the text base for each genre within a corpus. Distributions for these word frequencies were highly skewed, which is to be expected given that most SAT words appear very infrequently within the corpora, although some are quite common. Because the mean is strongly skewed under conditions of non-normality, we present the median as the primary measure of central tendency. Similarly, mean-based tests, such as the *t*-test, perform very poorly when distributions strongly deviate from normality. As a result, all analyses were conducted using nonparametric statistics. More specifically, we used the Mann-Whitney *U*-test, a rank-based method for testing whether values in one population are larger than another, to compare the frequency of SAT words in Fiction texts versus Nonfiction texts. Nonstandardized effect sizes are reported as the difference in medians, in light of the intuitively interpretable nature of our scores (i.e., differences in frequency, in units of words per million words). The R script and data to reproduce these analyses are available for download at osf.io/3rcsh.

Results

Findings for each corpus are presented in Table 1. In general, SAT words were found to appear less frequently in Fiction texts relative to Nonfiction texts. This is inconsistent with the hypothesis that word content differences can explain why exposure to fiction texts predicts verbal ability better than exposure to nonfiction. However, the results varied somewhat for each Nonfiction comparison group and some differences appeared across corpora. Importantly, none of these differences were

Table 1. Frequency of SAT Words (per Million Words) Within Fiction and Nonfiction Genres for Each Corpus.

Corpus	Genre	Median (MAD)	Min–Max	Skew	Kurtosis	*U*	*p*	Effect Size[a]	95% CI[b]
COCA									
	Fiction	0.82 (0.83)	0–39.67	7.73	83.55	—	—	—	—
	Academic	1.13 (1.29)	0–83.03	5.84	45.80	206,227.5	<.001	−0.31	[−0.50, −0.10]
	Magazines	1.04 (1.08)	0–135.26	14.21	244.20	202,753.5	<.001	−0.22	[−0.33, −0.11]
	Newspapers	0.91 (1.02)	0–173.47	17.15	340.33	183,945	.389	−0.09	[−0.21, 0.01]
BNC									
	Fiction	0.94 (1.02)	0–56.63	8.76	108.47	—	—	—	—
	Academic	0.98 (1.16)	0–88.71	5.82	49.85	187,156	.162	−0.04	[−0.26, 0.12]
	Magazines	0.83 (1.02)	0–106.44	12.76	193.80	182,657.5	.518	0.11	[0.01, 0.31]
	Newspapers	0.76 (0.86)	0–100.89	13.61	213.05	201,272	<.001	0.18	[−0.02, 0.27]
	Non-academic	1.03 (1.08)	0–152.77	13.91	250.97	189,342.5	.078	−0.09	[−0.25, 0.05]
Strathy									
	Fiction	0.80 (1.19)	0–69.76	10.43	151.46	—	—	—	—
	Academic	1.04 (1.13)	0–181.07	9.55	130.44	194,241.5	.010	−0.24	[−0.68, −0.08]
	Magazines	1.00 (1.19)	0–218.69	13.81	235.02	185,422.5	.267	−0.20	[−0.63, −0.11]
	Newspapers	0.46 (0.56)	0–355.61	17.30	329.05	213,555	<.001	0.34	[−0.03, 0.47]
	Nonfiction	1.19 (1.76)	0–139.39	12.62	198.14	193,718.5	.012	−0.39	[−0.78, −0.20]

MAD, median absolute deviation; CI, confidence interval.
[a]Effect sizes are the difference in medians between Fiction and the Nonfiction comparator. Negative effect sizes indicate greater frequency for the Nonfiction genre.
[b]Bootstrapped 95% confidence intervals of the effect size.

large in magnitude. The largest effect was a difference in frequency of 0.39 words per million (i.e., one extra appearance of a word every 2.56 million words). In the BNC and in Strathy, SAT words were found less frequently in the Nonfiction category of Newspapers compared with Fiction. For all three corpora, SAT words appeared with greater frequency in the nonfiction genre of Academic texts than in Fiction. Somewhat mixed results were observed for the category of Magazines across the three corpora. However, it is important to emphasize that all differences were small in magnitude. Inferential statistics, to which we now turn, largely confirmed these descriptive results.

For COCA, SAT words were found less often in Fiction (Mdn = 0.82) than either Academic journals (Mdn = 1.13; U = 206227.5, $p < .001$) or Magazines (Mdn = 1.04; U = 202753.5, $p < .001$). No statistically significant difference in the frequency of SAT words was observed between Fiction texts and Newspapers (Mdn = 0.91; p = .389).

Results obtained using the BNC were quite different from those found with COCA. In the BNC there were no statistically significant differences in SAT word frequency between Fiction texts (Mdn = 0.94) and either Academic journals (Mdn = 0.98; p = .162) or Magazines (Mdn = 0.83; p = .518). SAT words appeared more frequently in Fiction texts than in Newspapers (Mdn = 0.76; U = 201272, $p < .001$). Non-academic texts, a genre unique to the BNC, contained a greater frequency of SAT words than Fiction (Mdn = 1.03), although this difference just failed to attain threshold for statistical significance (p = .078).

For the Strathy, SAT words appeared less frequently in Fiction (Mdn = 0.80) compared with Academic journals (Mdn = 1.04; U = 194241.5, p = .010) but not Magazines (Mdn = 1.00; p = .267). As found in the BNC, SAT word frequency was higher in Fiction than in Newspapers (Mdn = 0.46; U = 213555, $p < .001$). Finally, SAT word frequency was higher in Nonfiction texts, a genre found only the in Strathy corpus, than in Fiction (Mdn = 1.19; U = 193718.5, p = .012).

In addition to investigating differences in the average frequency of SAT words, we also explored whether the distribution of moderately or highly frequent SAT words differed between fiction and nonfiction texts, based on a suggestion by an anonymous reviewer. Words appearing with non-zero levels of frequency may be more prevalent in fiction texts compared with nonfiction texts, and these words might be particularly relevant to examine because they are more likely to be encountered by readers. Figure 1 presents the number of words observed to appear with different levels of frequency for each genre across the three corpora. Moderately or highly frequent SAT words do not seem to appear more often in Fiction than in Nonfiction texts. If anything, these SAT words appear to be seen with slightly greater frequency in the nonfiction genres compared with fiction.

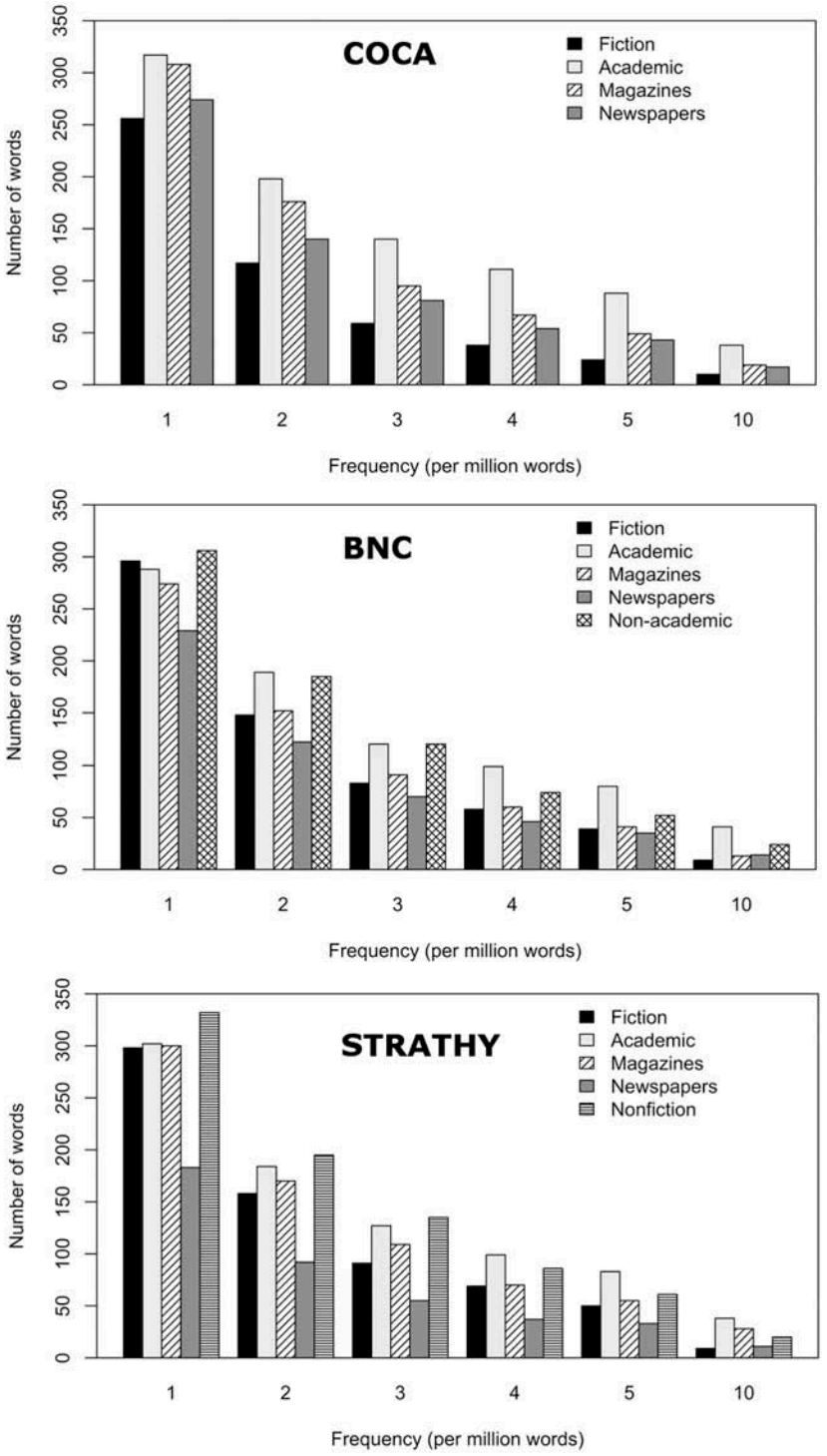

Figure 1. Histogram plots visualizing the number of words for each genre that appear with increasing levels of frequency across the corpora, beginning with ≥1 words per million up to ≥10 words per million, for the three different corpora: COCA, BNC, and Strathy.

The same reviewer noted that some SAT words did not appear at all in the texts sampled for each genre across the three corpora (see minimum word frequency values, Table 1). These zero values represent limitations of the corpora which, although they are large samples of texts, contain a minority of all possible texts. Within the true population of fiction texts, for example, it would be highly unlikely for any of these SAT words to never appear. If, however, the corpora are biased in the number of zero frequency words across genres (e.g., some words are more likely to fail to appear in the fiction texts compared to nonfiction texts, for some unusual reason not related to typical content or meaning), then these zero values may have biased our results. To rule out this possibility, we re-ran all our analyses changing these zero values to missing values. The results of this analysis largely replicated our main analysis. The only changes in statistical significance were found for the Strathy corpus, in which the median frequency for the Fiction category rose to 1.34 (median absolute deviation, 1.19) so that the difference between the Fiction and Academic categories (Mdn = 1.18; MAD = 1.23) became statistically nonsignificant (p = .10), whereas the difference between Fiction and Magazines (Mdn = 1.10; MAD = 1.19) became statistically significant, U = 152035, p = .004. However, it is worth keeping in mind that the effect sizes for these differences (and all differences reported) remain rather small.

Another potential concern, raised by this same reviewer, is that the maximum SAT word frequency values for Fiction appear to be far smaller than those for the other genre categories (Table 1). This might mean there is something unusual about those words that appear at a very high frequency in the categories that we chose to represent nonfiction. To systematically investigate this potential issue, we identified the highest frequency words in the Fiction category for each corpus and then found those words that appeared at an even higher frequency in the other genre categories. What we found was illuminating. The words that were more frequent than the most frequent Fiction word for COCA were INDIGENOUS, AUGUST, EMPIRICAL, SCOPE, DOCTRINE, COMPLIANCE, and INHERENT for the Academic category. For the Magazines and Newspapers categories, the same two words were more frequent: AUGUST and STOCK. For the BNC, the Academic category had two words more frequent than the highest frequency Fiction word, SCOPE and DOCTRINE. For the remaining categories (Magazines, Newspapers, and Non-academic), the same two words as before, AUGUST and STOCK, were more frequent. Finally, for Strathy, two words were identified across all Nonfiction categories, PROVINCIAL and AUGUST, with the word STOCK identified for Magazines and Newspapers. Based on these results, it appears that a few of these words, namely AUGUST, STOCK, and PROVINCIAL, are appearing with a high frequency in the Nonfiction genres as a result of their multiple potential meanings. Although the SAT may include these words based on their more literary meanings, these words are likely to appear at high frequencies in the other categories of text based on their other (and perhaps more banal) usages (e.g., the month August, "stock" in a financial sense, and "provincial" in a purely geographical sense). As a result, we repeated our analyses removing these three words. The results were very similar to our initial analysis, and no changes in statistical significance were observed across any of the corpora. Finally, combining the two approaches (i.e., changing zero values to missing values and removing the words AUGUST, STOCK, and PROVINCIAL) largely replicated what was found after removing the zero frequencies only. (Full details of these additional analyses are available upon request or reproducible based on the data and code made publicly available.)

Discussion

To examine whether word frequency differences can explain why narrative fiction is a better predictor of verbal ability than nonfiction, we examined the frequency of words typically tested in the verbal section of the SAT within Fiction and Nonfiction texts, based on three large corpora. Although SAT words were found to appear more frequently in Fiction texts than Newspapers for two corpora (the BNC and Strathy), the size of this effect (0.18–0.34 words per million) represents a

negligible and likely meaningless difference in SAT word exposure. To put the size of this effect in context, a difference of 0.34 appearances per million words means that a typical reader might see this word once more often every 30 books or so (assuming 100,000 words per book, a conservative estimate). In addition, SAT word frequencies were higher in most Nonfiction genres than in Fiction, although these effects were also very small in magnitude. Examining the number of words that appeared in each genre with different levels of frequency similarly failed to provide evidence that SAT words appear more often in fiction texts compared with nonfiction texts. As a result, there appears to be very little evidence that simple word frequency differences can account for why fiction is a better predictor of verbal ability than nonfiction. Interestingly, researchers from Japan have come to a similar conclusion based on a study of Japanese readers and texts. Inohara and Utsumi (2016) found that reading preferences for fiction predicted vocabulary, whereas a preference for nonfiction did not, consistent with past findings (Mar & Rain, 2015; McGeown et al., 2015; Spear-Swerling et al., 2010). Moreover, also very much in line with the results reported here, these researchers found that the content of the vocabulary test they used was best represented by the expository texts in their corpora and not the narrative texts, based on latent semantic analysis.

Some strengths of this study are the large number of SAT words examined and the investigation of three large corpora composed of texts from geographically distinct regions. For these reasons we believe our finding that there are no meaningful differences in SAT word frequency between fiction and nonfiction texts to be a robust and generalizable finding, not tied to a small number of possible words or English texts from a single country. Similar to the broad and diverse measurements of text exposure that led to the original discovery of fiction being a better predictor of verbal ability than nonfiction, the corpus analysis used in the present study permitted a broad and diverse sampling of relevant texts that are likely to represent fiction and nonfiction texts in general. That said, future studies should attempt to replicate these results using word lists derived from measures of verbal ability other than the SAT and/or using larger corpora. Regarding the latter, however, it is essential that the corpora include genre meta-data in order to investigate this question appropriately.

In light of the little evidence found that content differences between fiction and nonfiction texts could account for their differential relation to verbal ability, it would make the most sense to pursue other possible explanations for this effect. For example, it is possible that readers might be more motivated to read fiction closely compared with nonfiction (Logan, Medford, & Hughes, 2011). Greater intrinsic motivation to read fiction texts could arise because these texts are more personally relevant or more likely to elicit emotions relative to nonfiction (Mar, Oatley, Djikic, & Mullin, 2011). Close reading, in turn, could help facilitate language learning in readers through a deeper processing of the text than that observed during skimming or superficial reading. Along similar lines, higher intrinsic motivation to understand the events of a story compared with the arguments in an informational text could mean that readers are more likely to re-read sections of fiction compared with nonfiction, or more likely to look up words they do not understand. Another possible explanation could lie in the emotional nature of fiction texts. Stories seem more likely to elicit emotion than expository nonfiction (Oatley, 1994), with children expressing a strong emotional affinity for stories beginning at early ages (Alexander, Miller, & Hengst, 2001) and readers mentally representing and experiencing the emotions of story protagonists (Gernsbacher, Goldsmith, & Robertson, 1992; Laszlo & Cupchik, 1995; Mar et al., 2011). Emotion has long been known to facilitate memory (Hamann, 2001), with emotionally charged words better remembered than neutral words (LaBar & Phelps, 1998). If the emotions evoked by reading fiction help to imprint verbal content in memory, then perhaps this is the reason why verbal abilities are better predicted by exposure to fiction compared with nonfiction. One other possibility is that the rich representational context of narratives might make it easier to infer the meaning of unknown words in fiction compared with nonfiction (Nagy & Herman, 1987).

Alternative causal models and potential third-variable explanations should also be explored for us to best understand why fiction preferentially predicts verbal ability over nonfiction. For example, it is entirely possible that better verbal abilities motivate individuals to read more, which would in turn promote verbal ability. The excellent meta-analysis on reading and verbal

ability by Mol and Bus (2011) found evidence to support this sort of upward spiral. Specifically, they found that exposure to print predicts more and more variance in verbal abilities as you examine older and older populations, increasing from 12% in primary school students to 34% in college and university students. This is akin to the "rich-get-richer" or "Matthew effect" for reading, noted over 30 years ago (Stanovich, 1986; Walberg & Tsai, 1983). The key question to answer now, however, is why this would privilege fiction over nonfiction. For the Matthew effect to provide an explanation of the past results, verbal abilities would have to preferentially motivate people to read fiction over nonfiction. It is possible that although reading ability promotes reading in general, fiction is more often read as it is more intrinsically motivating to most readers, and this could then explain the differential association. Another important consideration is that some third variable might be responsible for promoting both verbal abilities and reading. Importantly, any causal third variable will need to be preferentially associated with fiction reading over nonfiction reading to account for the pattern of past results. For this reason some of the more obvious candidates, such as intellectual curiosity, do not appear on the surface to be promising explanations, as these would likely promote the reading of nonfiction and (or perhaps even more than) the reading of fiction. This is clearly an area that requires greater theoretical development and empirical attention. In past work, potential third-variable explanations have been explored by controlling for them statistically. These variables include gender, age, years of English fluency, biased responding, exposure to the other genre, and other reading skills (i.e., word reading and reading speed; Mar & Rain, 2015; McGeown et al., 2015). However, statistically controlling for confounds is a very difficult undertaking and researchers should interpret such results cautiously (Westfall & Yarkoni, 2016).

In closing, exploring whether and why reading fiction predicts verbal abilities better than reading nonfiction is a very promising topic for future investigation. In light of the findings of this current study, basic word content differences between fiction and nonfiction do not seem to be a likely explanation for this phenomenon. As a result, more attention should be directed at other possibilities, including alternative causal models and possible third-variable explanations. In addition, although fiction has been found to be a better predictor of verbal ability for adolescents (McGeown et al., 2015; Spear-Swerling et al., 2010) and adults (Mar & Rain, 2015), more research on this basic difference would help to better establish the generalizability and boundary conditions of this phenomenon.

References

Alexander, K. J., Miller, P. J., & Hengst, J. A. (2001). Young children's emotional attachments to stories. *Social Development*, *10*, 374–398.

Allen, L., Cipielewski, J., & Stanovich, K. E. (1992). Multiple indicators of children's reading habits and attitudes: Construct validity and cognitive correlates. *Journal of Educational Psychology*, *84*, 489–503.

Burt, J. S., & Fury, M. B. (2000). Spelling in adults: The role of reading skills and experience. *Reading and Writing*, *13*, 1–30.

Cipielewski, J., & Stanovich, K. E. (1992). Predicting growth in reading ability from children's exposure to print. *Journal of Experimental Child Psychology*, *54*, 74–89.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology*, *83*, 264–274.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*, 934–945.

Davidse, N. J., de Jong, M. T., Bus, A. G., Huijbregts, S. C. J., & Swaab, H. (2011). Cognitive and environmental predictors of early literacy skills. *Reading and Writing*, *24*, 395–412.

Davies, M. (2004). BYU-BNC: The British National Corpus [dataset]. Retrieved from http://corpus.byu.edu/bnc/

Davies, M. (2008). The Corpus of Contemporary American English: 520 million words, 1990–present [dataset]. Retrieved from http://corpus.byu.edu/coca/

Ecalle, J., & Magnan, A. (2008). Relations between print exposure and literacy skills: New evidence from grade 1–5. *British Journal of Developmental Psychology*, *26*, 525–544.

Gardner, D. (2004). Vocabulary input through extensive reading: A comparison of words found in children's narrative and expository reading materials. *Applied Linguistics*, *25*, 1–37.
Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. W. (1992). Do readers mentally represent characters' emotional states? *Cognition and Emotion*, *6*, 89–111.
Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*, 394–400.
Inohara, K., & Utsumi, A. (2016, August 10–13). Computational explanation of "fiction text effectivity" for vocabulary improvement: Corpus analyses using latent semantic analysis. Poster presented at the 38th Annual Meeting of the Cognitive Science Society, Philadelphia, PA.
LaBar, K. S., & Phelps, E. A. (1998). Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science*, *9*, 490–493.
Laszlo, J., & Cupchik, G. C. (1995). The role of affective processes in reading time and time experience during literary reception. *Empirical Studies of the Arts*, *13*, 25–37.
Logan, S., Medford, E., & Hughes, N. (2011). The importance of intrinsic motivation for high and low ability readers' reading comprehension performance. *Learning and Individual Differences*, *21*, 124–128.
Mar, R. A., Oatley, K., Djikic, M., & Mullin, J. (2011). Emotion and narrative fiction: Interactive influences before, during, and after reading. *Cognition and Emotion*, *25*, 818–833.
Mar, R. A., & Rain, M. (2015). Narrative fiction and expository nonfiction differentially predict verbal ability. *Scientific Studies of Reading*, *19*, 419–433.
Martin-Chang, S. L., & Gould, O. N. (2008). Revisiting print exposure: Exploring differential links to vocabulary, comprehension and reading rate. *Journal of Research in Reading*, *31*, 273–284.
McGeown, S. P., Duncan, L. G., Griffiths, Y. M., & Stothard, S. E. (2015). Exploring the relationship between adolescents' reading skills, reading motivation and reading habits. *Reading and Writing*, *28*, 545–569.
Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, *137*, 267–296.
Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M.G. McKeown & M.E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19–36). Mahwah, NJ: Lawrence Erlbaum Associates.
Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge, UK: Cambridge University Press.
Oatley, K. (1994). A taxonomy of the emotions of literary response and a theory of identification in fictional narrative. *Poetics*, *23*, 53–74.
Sparks, R. L., Patton, J., & Murdoch, A. (2014). Early reading success and its relationship to reading achievement and reading volume: Replication of "10 years later". *Reading and Writing*, *27*, 189–211.
Spear-Swerling, L., Brucker, P. O., & Alfano, M. P. (2010). Relationships between sixth-graders' reading comprehension and two different measures of print exposure. *Reading and Writing*, *23*, 73–96.
Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407.
Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory and Cognition*, *20*, 51–68.
Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*, 402–433.
Strathy Language Unit. (2013). The Strathy Corpus of Canadian English [dataset]. Retrieved from http://corpus.byu.edu/can/
Walberg, H. J., & Tsai, S. (1983). Matthew effects in education. *American Educational Research Journal*, *20*, 359–373.
West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its correlates. *Reading Research Quarterly*, *28*, 35–50.
Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PLoS One*, *11*, e0152719.