



# Using the DRM paradigm to assess language processing in monolinguals and bilinguals

Ellen Bialystok<sup>1</sup> · Avanti Dey<sup>2</sup> · Margot D. Sullivan<sup>3</sup> · Mitchell S. Sommers<sup>4</sup>

Published online: 23 January 2020  
© The Psychonomic Society, Inc. 2020

## Abstract

Both languages are jointly activated in the bilingual brain, requiring bilinguals to select the target language while avoiding interference from the unwanted language. This cross-language interference is similar to the within-language interference created by the Deese–Roediger–McDermott false memory paradigm (DRM; Roediger & McDermott, 1995, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21[4], 803–814). Although the mechanisms mediating false memory in the DRM paradigm remain an area of investigation, two of the more prominent theories—implicit associative response (IAR) and fuzzy trace—provide frameworks for using the DRM paradigm to advance our understanding of bilingual language processing. Three studies are reported comparing accuracy of monolingual and bilingual participants on different versions of the DRM. Study 1 presented lists of phonological associates and found that bilinguals showed higher rates of false recognition than did monolinguals. Study 2 used the standard semantic variant of the task and found that bilinguals showed lower false recognition rates than did monolinguals. Study 3 replicated and extended the findings in Experiment 2 in another semantic version of the task presented to younger and older adult monolingual and bilingual participants. These results are discussed within the frameworks of IAR and fuzzy-trace theories as further explicating differences between monolingual and bilingual processing.

**Keywords** Bilingualism · Selective attention · False memory · Semantics · Phonology · Aging

It is well-documented that bilinguals have continual access to information from both languages during language use, even in strongly monolingual contexts (De Groot, 1992; Dijkstra & Van Heuven, 1998; Francis, 1999; Green, 1993; Kroll, 1993; Kroll & Stewart, 1994; Spivey & Marian, 1999; for review see Kroll, Dussias, Bice, & Perrotti, 2015). This parallel activation is found for phonological and semantic features in both the auditory (Brysbaert, Van Dyck, & Van de Poel, 1999; Dijkstra, Grainger, & Van Heuven, 1999; Ju & Luce, 2004; Kaushanskaya & Marian, 2007; Linck, Kroll, & Sunderman,

2009; Schulpen, Dijkstra, Schriefers, & Hasper, 2003; Spivey & Marian, 1999; Weber & Cutler, 2004; Zhao & Li, 2010) and visual modality (Dijkstra, Van Heuven, & Grainger, 1998; Grainger, 1993), as well as in language production (Colomé, 2001; Costa, Miozzo, & Caramazza, 1999; Hermans, Bongaerts, De Bot, & Schreuder, 1998; Kroll, Bobb, & Wodniecka, 2006), and extends to speech–sign bilingualism (Morford, Wilkinson, Villwock, Piñar, & Kroll, 2011).

Two demonstrations of this joint activation are particularly compelling. First, Marian and Spivey (2003a, 2003b) presented a visual world task to Russian–English bilinguals in which participants were asked to make eye movements to a named picture. The context, instructions, and environment were entirely in English, and no mention was made of Russian. Nonetheless, lure items in the visual world display that had phonological similarity (e.g., “marka”, meaning “stamp”) to the target item (“marker”) elicited significant eye movements from the bilinguals. Second, Thierry and Wu (2007) asked Chinese–English bilinguals at an English-speaking university to judge the semantic relatedness of English word pairs while ERP was recorded. The results showed that English word pairs that were not semantically related but shared Chinese characters if they would be translated into Chinese and then

✉ Ellen Bialystok  
ellenb@yorku.ca

<sup>1</sup> Department of Psychology, York University, 4700 Keele St., Toronto, ON M3J 1P3, Canada

<sup>2</sup> Center for Vital Longevity, University of Texas at Dallas, Richardson, TX, USA

<sup>3</sup> Department of Psychology, Ryerson University, Toronto, ON, Canada

<sup>4</sup> Department of Psychological and Brain Sciences, Washington University in St. Louis, St. Louis, MO, USA

written elicited reduced N400 amplitude, as is found for semantically related pairs. Thus, judgments of the English words were unconsciously influenced by the written forms of their Chinese translations.

One consequence of jointly activated languages is that lexical access is more difficult for bilinguals than for monolinguals, as is shown by performance in picture naming (Friesen, Chung-Fat-Yim, & Bialystok, 2016; Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Sullivan, Poarch, & Bialystok, 2018), tip-of-the tongue events (Gollan & Silverberg, 2001), and verbal fluency tasks (Giezen & Emmorey, 2017; Gollan, Montoya, & Werner, 2002; Luo, Luk, & Bialystok, 2010). Yet, despite both languages being constantly active, bilinguals rarely make language intrusion errors (Gollan, Sandoval, & Salmon, 2011; Myers-Scotton, 2002; Sandoval, Gollan, Ferreira, & Salmon, 2010). An alternative account in which bilinguals establish fewer automatic links with each language because of lower frequency of use for each language, called the weaker links view (Gollan, Montoya, Cera, & Sandoval, 2008), is consistent with the absence of intrusion errors, but provides a less clear account of effortful retrieval than does joint activation (Sullivan et al., 2018). Therefore, the joint activation account requires that bilinguals have developed efficient processes for language selection. This facility with language selection not only changes how language processing is carried out in bilinguals but may also be partly responsible for evidence that bilinguals frequently outperform monolinguals on nonverbal tasks based on conflict and selection (see a review in Bialystok, 2017). However, this research has produced inconsistent results, particularly with young adults, where some studies report significant benefits for bilinguals on these tasks (e.g., Bialystok, Craik, Klein, & Viswanathan, 2004; Costa, Hernandez, & Sebastian-Galles, 2008) and others fail to find such outcomes (e.g., Paap & Greenberg, 2013; von Bastian, Souza, & Gade, 2016). Thus, identifying the mechanisms involved in bilingual language selection takes on greater consequence since understanding bilingual language processing may have implications for models of cognitive processing more broadly.

If bilinguals need to constantly deal with jointly activated representations that interfere with the present task, then at some level ordinary language selection for bilinguals resembles the configuration created in the Deese–Roediger–McDermott (DRM) false memory paradigm (Roediger & McDermott, 1995). In the original DRM paradigm (Roediger & McDermott, 1995), participants are exposed to a list of words (e.g., *thread, pin, sewing, sharp, point*) that are all semantically associated with a word that is not presented (*needle*), and then asked to recognize or recall the original words. The typical finding is that despite being warned against guessing, participants falsely recall and recognize the nonpresented associate, termed a *critical lure*, at rates that equal or exceed items that were actually presented (Payne,

Elie, Blackwell, & Neuschatz, 1996; Roediger, McDermott, Pisoni, & Gallo, 2004; Roediger, McDermott, & Robinson, 1998; Schacter, Norman, & Koutstaal, 1998).

In a related study, Sommers and Lewis (1999) examined whether lists of phonological associates to a nonpresented critical lure would produce rates of false recall and recognition similar to those obtained with lists of semantic associates (Roediger & McDermott, 1995). Analogous to the original DRM paradigm, Sommers and Lewis presented participants with lists composed of phonological neighbors (words differing from a target word by the addition, deletion or substitution of a single phoneme) of a nonpresented critical lure. For example, one list contained phonological neighbors of the critical lure *cat* (*kit, cab, hat, bat . . .*), although *cat* was not presented. As in the DRM paradigm with semantic associates, Sommers and Lewis reported that participants remembered critical lure items at rates equal to or higher than those for list items.

### Mechanisms mediating false memories in the DRM paradigm and implications for bilingual language processing

The two most prominent theoretical accounts of false memories in the DRM paradigm are implicit associative response (IAR) and fuzzy-trace theories (see Roediger et al., 1998, for review). According to IAR, presentation of list items increases activation on representations of those items as well as on semantic associates, as activation spreads through semantic networks (Collins & Loftus, 1975). Thus, presentation of each item in a list of semantic associates will activate not only the item itself but also the semantically related critical lure. Similarly, within activation-competition models of word recognition (Luce & Pisoni, 1998; Norris, Cutler, McQueen, & Butterfield, 2006), presentation of the target item increases activation on representations of that word and then spreads to phonologically similar words. If spreading activation from list items to critical lures is sufficiently high, participants will falsely recall or recognize the critical lure just like any other list item. Thus, within IAR, both semantic and phonological false memories occur because participants mistakenly ascribe high activation levels on the critical lure to it having been presented.

The IAR theory of false memories places the locus of false recall and recognition at encoding; activation of list items during encoding spreads automatically to related items, and false memories occur when levels of activation are sufficiently high on these related, but nonpresented, critical lures. A variant of IAR theory, activation-monitoring theory, is similar in that it stresses the role of spreading activation from list items to nonpresented critical lures, but places the locus of false memories at retrieval, rather than at encoding. According to

an activation-monitoring account, false memories result primarily from source monitoring failures; both phonological and semantic false memories occur when individuals misattribute activation on critical lures to actual item presentation rather than from spreading activation from presented items (Finley, Sungkhasettee, Roediger, & Balota, 2017). That is, false memories arise according to activation-monitoring accounts because of errors in ascribing the source of activation on lexical representations. In the current work, we focus on the more general IAR theory because it presents a situation more analogous to language selection in bilinguals, which focuses on coactivation within the two lexicons rather than on source monitoring processes, and which is therefore more consistent with the goal of using theories of DRM to derive predictions about language processing in monolinguals and bilinguals rather than to support specific theories of false memories in the DRM paradigm.

According to the fuzzy-trace account of false memories (Brainerd & Reyna, 2002), two different kinds of representations are created during encoding word lists: gist traces, which contain the general thematic meaning of each list but not item-specific perceptual details and verbatim traces which contain item-specific details. Within the fuzzy-trace account, false memories occur when participants rely on gist rather than verbatim representations for recall and recognition. Although fuzzy-trace accounts of false memories have been largely restricted to lists of semantic associates, as the theory stresses formation of gist representations based on meaning, a recent study (McGeown, Gray, Robinson, & Dewhurst, 2014) used the fuzzy-trace framework to explain the relationship between language skills and susceptibility to phonological false memories. McGeown et al. (2014) found that phonological awareness, a measure of individuals' ability to recognize relationships between phonemes, was related to susceptibility to phonological false memories. Thus, although fuzzy-trace accounts of false memories are more developed for lists of semantic associates, there is no reason that the approach cannot be extended to lists of phonological associates.

In considering the different accounts of DRM false memories, it is important to note that the purpose of the current study is not to adjudicate between IAR and fuzzy-trace explanations of false memories in the DRM paradigm but rather to use the DRM paradigm as a tool to understand differences between monolingual and bilingual language processing. The accounts proposed by the two theories provide useful frameworks for generating predictions about differential susceptibility of monolinguals and bilinguals to DRM false memories, and, as we note below, in some cases the two theories make complementary predictions, and in others, contrasting predictions.

Consider first predictions based on fuzzy-trace theory. As noted, within fuzzy-trace theory, false memories arise from reliance on gist, rather than on verbatim representations

created by presenting semantically or phonologically related word lists. Differences in the extent to which particular groups rely on verbatim versus gist representations have been used to account for developmental differences in susceptibility to false memories in children, young adults, and older adults, as well in individuals with dyslexia (Brainerd & Reyna, 2002; Gomes, Cohen, Desai, Brainerd, & Reyna, 2014; Holliday, Brainerd, & Reyna, 2011; Obidziński & Nieznański, 2017). To our knowledge, there has not been a direct comparison of bilinguals' and monolinguals' reliance on gist versus verbatim representations. However, if bilinguals have increased reliance on gist, rather than verbatim representations, we would predict greater levels of false recognition in this group than in monolinguals. Conversely, if bilinguals rely less on gist representations than do monolinguals, we would expect lower levels of both phonological and semantic false memories. No differences between the groups would indicate similar reliance on the two types of representations. Thus, the findings from the current study should provide the first, albeit indirect, evidence regarding relative reliance on gist versus verbatim representations in monolinguals and bilinguals.

According to IAR, differential rates of false memories in monolinguals and bilinguals could arise from two (not mutually exclusive) processes. First, according to IAR, activation spreads from presented items to phonological or semantic associates including the critical lure, creating joint activation of presented and nonpresented items. Differences in either the initial levels of activation and/or the degree to which activation spreads throughout semantic and phonological networks would predict differential activation of phonological or semantic associates for bilingual and monolingual individuals and hence differences in susceptibility to false memories. A second mechanism that could account for differential susceptibility of monolinguals and bilinguals to false memories according to IAR is differences in attentional control systems. Participants in the DRM paradigm must selectively report activated target items and avoid recalling or recognizing the activated, but nonpresented associated items. If both language processing in bilinguals and word identification in the DRM take place in the context of jointly activated alternatives that require selective attention to evaluate, then the experience of bilinguals may improve their ability to avoid critical lures in the DRM because it relies on processes similar to those used in bilingual language selection. To that end, the general prediction is that bilinguals will be less susceptible than monolinguals to false alarms in the DRM task. To the extent that both bilingual language processing and DRM recruit similar processes of selective attention, the more practiced and therefore more automatic attentional processes of bilinguals should benefit their performance on DRM, as reflected in reduced levels of false recognition.

## Semantic versus phonological false memories

Although the majority of research with the DRM paradigm has used lists of semantic associates, parallel findings have also been observed using lists of phonological associates (Ballou & Sommers, 2008; Sommers & Lewis, 1999; Wallace, Stewart, & Malone, 1995; Watson, Balota, & Sergent-Marshall, 2001; Westbury, Buchanan, & Brown, 1999). Interestingly, investigations examining the mechanisms mediating these two types of false memory suggest that they may be generated by distinct processes (Ballou & Sommers, 2008; Holliday & Weekes, 2006; Watson et al., 2001). For example, Ballou and Sommers (2008) found no correlation between susceptibility to semantic and phonological false memories in a group of young adults. Holliday and Weekes (2006) reported different developmental trajectories for phonological and semantic false memories in children ages 8–13 years, with false memories increasing with age for lists of semantic associates and decreasing with age for lists of phonological associates. Therefore, one goal of the present study was to compare susceptibility to false memories in monolinguals and bilinguals for lists of both semantic and phonological associates. Based on findings suggesting differences between the two types of false memory, it may be that bilingualism has different effects on phonological versus semantic false memory.

In summary, in the absence of evidence to suggest differential reliance on verbatim versus gist representations, fuzzy-trace theory would predict no differences in susceptibility to DRM false memories. Evidence contrary to this hypothesis would provide the first findings to suggest that bilingualism may be associated with differential reliance on gist versus verbatim representations. In contrast, according to IAR, differential susceptibility to DRM false memories can result from differences in the magnitude of spreading activation as well as differences in selection of activated items. To the extent that bilinguals have improved attentional control relative to monolinguals, we predict reduced susceptibility to DRM false memories in this group. Furthermore, based on evidence for a dissociation between phonological and semantic false memories, it may be that that bilingualism has differential effects on these two types of false memories.

The three experiments reported in the current study examined rates of false memory in monolingual and bilingual participants across different types of linguistic features and different populations. Experiment 1 presented a phonological version of the task to young adults; Experiment 2 used the standard semantic version with young adults; and Experiment 3 extended the design to compare younger and older adult performance. In all three experiments, we compared the performance of monolinguals and bilinguals.

## Experiment 1

### Method

#### Participants

Forty-five English-native monolingual young adults and 40 bilingual young adults participated in the study. This sample size was based on findings from Sommers and Lewis (1999), who found levels of phonological false memories comparable to those reported for semantic false memories (Roediger & McDermott, 1995) using a sample size of 42 young adults. All participants were undergraduate students at Washington University in St. Louis and were compensated with course credit. Bilingual participants reported being fluent in English and one of five other languages, including Cantonese, French, German, Mandarin, and Spanish. Bilingual participants typically started learning English concurrent with or shortly after learning each of these primary languages, but always before puberty. Thus, the bilingual participants were either simultaneous or non-English-dominant bilinguals.

#### Tasks

**Background measures** Participants were administered the Wechsler Adult Intelligence Scale (WAIS-III) Vocabulary subtest as a test of English vocabulary and the Cattell Culture Fair Intelligence Test (Cattell & Cattell, 1960) as a test of nonverbal reasoning. Participants were also administered the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007) to assess the degree to which speakers rated their use and fluency in all languages.

**DRM task** The auditory stimuli were selected from the lists generated by Sommers and Lewis (1999) for the phonological DRM task. These word lists of phonological associates were designed to parallel those used by Roediger and McDermott (1995) in their demonstrations of false recall and recognition with semantic associates. Each participant heard eight words lists, each of which contained 15 phonological associates of a (nonpresented) critical lure. A male talker with a Midwestern dialect recorded the words in a double-walled sound-attenuating booth. The productions were transduced using a free-field microphone, low-pass filtered at 8.5 kHz, and digitized online using a 16-bit A/D converter and a 22-kHz sampling rate. The root-mean squared (RMS) amplitude level of all words was digitally equated. Participants were seated in front of a PC and keyboard. The DRM task was run on SuperLab (Version 2.0.4; Cedrus Corporation). Stimuli were presented binaurally through headphones (Sennheiser HD 265).

**Study phase** The eight lists of 15 words each resulted in 120 unique study items. The order of list presentation was randomized across participants, but the order of the words within each list was pseudorandomized and remained constant for all participants. The beginning of each list was indicated by a 1,000-ms 500-Hz tone. Within each list, there was an interstimulus interval (ISI) of 2,000 ms. There was no response required on the part of the participant, although to ensure that sufficient attention was being engaged, participants were required to press the space bar between each list (i.e., after each tone). Participants were told to remember the words for a subsequent memory test.

**Recognition phase** The recognition test consisted of 48 items, including one critical lure related to each study list (8), two target items that had been presented on each of the lists (16), and 24 new items. New items were taken from nonpresented lists (see Sommers & Lewis, 1999, for a complete set of studied and new items). One item was selected randomly from each of the 16 nonpresented lists and then eight of the lists were selected randomly to obtain one additional item (i.e., one new item was taken from each nonpresented list with an additional item taken from eight of the lists). Thus, the test included 16 studied items and 32 new items, of which eight were critical lures. Words were presented one at a time in random order, and each was followed by a visual screen display of ‘OLD or NEW?’ Participants were asked to press one of two keyboard keys to indicate whether they recognized the word from the study phase. Participants were given up to 5,000 ms to respond to each word, after which the next word appeared.

## Results and discussion

### Background measures

Table 1 shows age and mean background measure scores for each language group. Monolinguals and bilinguals were comparable on age,  $t(83) = 1.11, p = .27$ , English vocabulary,  $t(83) = 1.66, p = 0.10$ , and monolinguals had marginally higher nonverbal intelligence than did bilinguals,  $t(83) = 1.77, p = 0.08$ . Monolinguals rated themselves as significantly more proficient in English than did bilinguals,  $t(83) = 3.83, p < .001, d = 1.12$ , and bilinguals rated themselves as significantly more proficient in their non-English language than in English,  $t(83) = 13.62, p < .001, d = 4.38$ . Monolingual speakers used English significantly more often than did bilinguals,  $t(83) = 12.61, p < .001, d = 4.06$ .

### DRM task

Table 2 shows the mean percentage of ‘OLD’ responses for studied, unstudied, and critical lure items. Separate one-way

ANOVAs were conducted to compare the performance of monolinguals and bilinguals on studied items, nonstudied items (other than critical lures), and critical lures. The accuracy of detecting the presented words was examined in a  $d$ -prime ( $d'$ ) analysis comparing correct responses to studied words against false alarms (excluding critical lures). These values, reported in Table 2, did not differ across groups,  $F < 1$ . The two language groups also did not differ on either correct recognition of studied items or false alarms to nonstudied items other than the critical lures (all  $F$ s  $< 1$ ). In contrast, bilinguals gave significantly more “old” responses to critical lures than did monolinguals  $F(1, 83) = 6.30, p = .01, \eta_p^2 = .07$ .

A sensitivity analysis was conducted in G\*Power to determine the minimum effect size that the design could reliably detect using power = .80 and  $\alpha = .05$ . The results indicated that a reliable group difference for response to the critical lures required minimally a critical  $F$  value of 3.95, with Cohen’s  $d$  of 0.24. The actual values obtained were  $F = 6.30$  and Cohen’s  $d = 0.55$ , both well beyond this limit.

As indicated in the analysis of demographic measures shown in Table 1, bilinguals had lower self-rated English-language proficiency than did monolinguals. It is possible that lower English proficiency could have contributed to the increased incidence of DRM false memories for bilinguals. To examine whether lower levels of English proficiency were associated with higher levels of false recognition, we correlated these measures for the bilinguals (monolinguals were near ceiling for ratings of English proficiency). English language proficiency was not significantly correlated with the number of phonological critical lures individuals recognized,  $r = -.16, p > .3$ .

Bilinguals produced higher rates of false recognition of critical lures than did monolinguals, despite comparable performance on studied items and new items other than critical lures. Based on IAR and fuzzy-trace theories, we initially predicted either no difference in false memories between monolinguals and bilinguals (fuzzy-trace theory) or reduced false memory for bilinguals (IAR). However, there was a significant difference between groups, but it was in the opposite direction than we had predicted; bilinguals were more susceptible than monolinguals to phonological false memories.

Differences in processing for phonological and semantic features is consistent with models that distinguish between operations that support word form and those that support word meaning (Potter, So, Von Eckardt, & Feldman, 1984; Snodgrass, 1984). Moreover, it may be that rates of semantic and phonological false memories in the DRM paradigm are mediated by distinct mechanisms (Ballou & Sommers, 2008; Chan, McDermott, Watson, & Gallo, 2005; Watson, Balota, & Roediger, 2003). Therefore, Experiment 2 tested a new group of monolingual and bilingual young adults using a semantic version of the DRM task.

**Table 1** Means (and SDs) of background variables by language group in Studies 1 and 2

Experiment	Study 1		Study 2	
	Monolingual	Bilingual	Monolingual	Bilingual
<i>n</i>	45	40	25	25
Age (years)	19.9 (2.2)	19.3 (1.2)	20.2 (2.7)	19.8 (0.9)
English vocabulary (out of 66)	54.0 (8.5)	49.1 (7.9)	52.5 (5.6)	50.1 (5.2)
Nonverbal intelligence	114.0 (7.4)	109.7 (5.8)	114.2 (8.5)	115.9 (13.6)
<i>LEAP-Q self-ratings<sup>a</sup></i>				
Age learned English (years)	1.4 (0.6)	5.8 (1.6)	1.6 (0.6)	6.5 (3.2)
Age learned non-English language (years)	—	1.8 (0.7)	—	2.0 (1.1)
Self-rated English Proficiency (out of 10)	9.5 (0.5)	7.5 (2.4)	9.7 (0.4)	7.6 (1.9)
Self-rated non-English proficiency (out of 10)	2.1 (1.4)	8.2 (1.3)	1.4 (0.6)	8.3 (1.3)

<sup>a</sup> The LEAP-Q contains 23 items, but we included only those that were most relevant to the current experiments. Self-report ratings of proficiency range from 0 = “no proficiency” and 100 = “fully fluent” and usage from 0 = “all other language” and 100 = “all English”. Proficiency ratings were based on an average of speaking and understanding. Usage ratings were based on an average consisting of ratings for three usage contexts

## Experiment 2

### Method

#### Participants

New groups of bilingual and monolingual participants from the same university as in Experiment 1 were recruited to perform the semantic version of the DRM. Twenty-five English-native monolingual and 25 bilingual young adults were recruited for Study 2. Bilingual participants were fluent in English and either Cantonese, German, Mandarin, or Spanish.

#### Tasks

**Background measures** Participants were administered the same tasks for English vocabulary (WAIS-III), nonverbal intelligence (Cattell Culture Fair Intelligence Test), and self-report measures from the LEAP-Q as were used in Study 1. Mean scores are reported in Table 1.

**DRM task** Eight lists generated by Roediger and McDermott (1995) were selected. To match the task design as closely as possible to that in Study 1. Each list contained 15 semantic associates of a critical lure. Mean backward association strength (BAS), the degree of association from studied items to the critical lure, was  $M = .23$  ( $SD = .05$ ). The recognition test included 48 items consisting of eight critical lures from the studied lists, two items from each of the eight studied lists, and 24 new items. New items were selected randomly from nonpresented lists as in Experiment 1. The procedure was identical to Study 1.

### Results and discussion

#### Background measures

As in Study 1, monolinguals and bilinguals were comparable on age,  $t(48) = 1.52, p = .32$ , English vocabulary,  $t(48) = 1.36, p = 0.17$ , and nonverbal intelligence,  $t(48) = 0.7, p = 0.48$ . Monolinguals rated themselves significantly more proficient in English than did bilinguals,  $t(48) = 4.83, p < .001, d = 1.51$ ,

**Table 2** Mean percentage of ‘OLD’ responses (and standard error of the mean) for studied, unstudied, critical lure items, and *d*-prime by language group in Studies 1 and 2

Experiment	Study 1: Phonological		Study 2: Semantic	
	Monolingual	Bilingual	Monolingual	Bilingual
Studied words	66.02 (2.63)	65.31 (2.80)	64.00 (2.81)	67.33 (2.81)
Unstudied words	23.33 (2.74)	22.82 (2.18)	24.00 (2.91)	23.00 (2.91)
Critical lures	54.17 (3.04)	65.31 (3.22)	59.52 (1.63)	47.51 (2.53)
<i>D</i> -prime <sup>a</sup>	1.18 (0.10)	1.23 (0.11)	1.15 (0.13)	1.23 (0.10)

<sup>a</sup> *D*-prime was calculated as hits (correct OLD response to studied items) minus false alarms (incorrect OLD response to new items) excluding critical lures

while bilinguals rated themselves as significantly more proficient in a non-English language than did monolinguals,  $t(48) = 21.09$ ,  $p < .001$ ,  $d = 7.16$ . Monolingual speakers used English significantly more often than did bilinguals,  $t(48) = 13.67$ ,  $p < .001$ ,  $d = 4.33$ , while bilinguals had significantly more non-English language exposure than did monolinguals,  $t(48) = 45.81$ ,  $p < .001$ ,  $d = 14.46$ .

### DRM task

Table 2 shows the mean percentage of ‘OLD’ responses for studied, unstudied, and critical lure items. *D*-prime analyses for the ability to detect old items are reported in Table 2 and indicated no difference between groups,  $F < 1$ . Separate one-way ANOVAs conducted on percentage of “old” responses to the three types of stimuli indicated no significant differences between monolinguals and bilinguals for correct recognition of studied items, and no significant difference for false alarms to new items that were not critical lures (all  $F$ s  $< 1$ ). Monolinguals, however, had significantly more “old” responses to critical lures than did bilinguals  $F(1, 48) = 5.80$ ,  $p = .01$ ,  $\eta_p^2 = .11$ . Using the same parameters as in Experiment 1 to determine the reliability of this finding, a sensitivity analysis indicated the need for a critical  $F$  value of 4.04 and a Cohen’s  $d$  of 0.32. The present results produced an  $F$  value of 5.80 and Cohen’s  $d$  of 0.69, again exceeding the necessary threshold. As in Experiment 1, monolinguals had higher self-rated English proficiency than did bilinguals. Also, as in Experiment 1, the self-rated English proficiency was not significantly related to the number of semantic false memories,  $r = .26$ ,  $p > .2$ .

In contrast to the results of Study 1, in which bilinguals showed more false alarms to phonologically related word lists than monolinguals did, bilinguals in Study 2 showed fewer false alarms to semantically related critical lures than did monolinguals. In both studies, the word lists were presented auditorily, ruling out presentation modality as a contributing factor. Instead, the findings from Study 2 are consistent with the initial prediction from IAR that bilinguals were less susceptible to critical lures than were monolinguals. Moreover, responses to the other two types of stimuli (studied words and nonstudied items other than the critical lure) were consistent across the two studies, with about 65% correct recognition of target items and about 23% false alarms to new noncritical items incorrectly called ‘OLD’. It was only the response to the critical lures that changed across studies and across language groups. Differences between monolinguals and bilinguals in their responses to critical lures was not the result of differential experience with English; the two groups were equally capable of correctly remembering target words and had equal false alarm rates for new items that were not critical lures.

Experiment 3 extended these findings by administering a different version of a semantic DRM that manipulated

backward associative strength (BAS) in the word lists and used written stimuli instead of oral presentation. The study also included older adults to determine if there are effects of aging. Experiment 3 was conducted independently of the first two experiments in a different location; neither group of researchers was aware of the ongoing study by the other group. Therefore, the procedures and background measures were somewhat different for the two research groups, but the possibility of converging results was therefore more compelling.

### Experiment 3

DRM studies of aging have reliably found increases in false memory rates for both healthy (Balota et al., 1999; Dehon & Brédart, 2004; Dennis, Kim, & Cabeza, 2007; Norman & Schacter, 1997; Tun, Wingfield, Rosen, & Blanchard, 1998) and patient populations (Balota et al., 1999; Budson, Sullivan, Daffner, & Schacter, 2003; Sommers & Huff, 2003). Activation-based accounts suggest that age-related increases in false memories result from reductions in attentional control (Balota et al., 1999; Gallo & Roediger, 2003; Watson et al., 2003). Because aging is associated with reduced attentional control (Gazzaley, Cooney, Rissman, & D’esposito, 2005; Hasher & Zacks, 1988; McDowd & Filion, 1992; see Verhaeghen & Cerella, 2002, for review), older adults are more impaired than are young adults in their ability to reduce activation on associated but nonpresented items, including the critical lures (Balota et al., 1999; Watson et al., 2001). However, there is also evidence that older bilinguals continue to demonstrate better attentional control than do older monolinguals (Bialystok, 2017, for review), and consequently there may be greater age-related increases in DRM false memories for monolinguals than for bilinguals. In contrast, from a fuzzy-trace perspective, age-related increases in DRM false memories result from an increased reliance on gist, rather than on verbatim representations with age (Brainerd & Reyna, 2002). Therefore, fuzzy-trace theory would predict that both older monolinguals and bilinguals should be more susceptible to semantic false memories than their younger counterparts, but language status (bilingual versus monolingual) should not matter.

In addition to testing young and older adults, Study 3 also included lists that varied in BAS to manipulate the strength of association between studied items and the critical lure. Lists with high BAS to the critical lure typically produce higher levels of false memories because the initial activation and/or the magnitude of spreading activation from list items to the critical lure is higher than in low BAS lists (Gallo & Roediger, 2002; Howe, Wimmer, & Blease, 2009). To the extent that bilingual older adults have better preserved attentional control than monolinguals, we expect bigger differences between these two groups for lists with high compared with low

BAS, particularly for older adults. That is, because selective attention demands are increased for lists with high BAS—owing to increased activation on nonpresented items—the better preserved attentional control of bilinguals might be particularly useful for resisting false memories and might magnify any differences between bilingual and monolingual older adults.

**Method**

**Participants**

A total of 129 participants were recruited for Experiment 3, and the final sample size was 112, divided as follows: 59 young adults recruited from the York University undergraduate psychology participant pool, consisting of 26 monolingual English speakers and 33 bilinguals who reported being fluent in English and at least one additional language out of 20 different languages; 53 healthy older adults recruited from the University of Toronto and York University participant pools, consisting of 26 monolingual speakers of English and 27 bilinguals who reported being fluent in English and at least one additional language out of 19 languages. Seventeen recruited participants were excluded from analyses due to low vocabulary or nonverbal intelligence (standard score <70, i.e., <2 SDs, *n* = 8; one younger monolingual, seven younger bilinguals), unclear language background (*n* = 7; four older adults, three younger adults), computer error (*n* = 1, younger monolingual), or visual impairment (*n* = 1, older bilingual). Table 3 reports participant background information.

**Tasks**

**Language and Social Background Questionnaire (LSBQ;** Anderson, Mak, Keyvani Chahi, & Bialystok, 2018). Participants answered questions pertaining to age, sex, handedness, education, vision/hearing problems, country of birth, as well as language use and self-rated fluency (relative to a native speaker) for all known languages (rated on scales of 0 to 100, where 0 represents *no proficiency* and 100 represents *fully fluent*).

**English receptive vocabulary measures** Participants were administered either the verbal component of the Shipley-2 (Shipley, Gruber, Martin, & Klein, 2009) or the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997) as a measure of English receptive vocabulary. Each test was administered according to standardized instructions and converted to normed scores by a set of tables based on age with a population mean of 100 and a standard deviation of 15.

**Nonverbal (fluid) intelligence measures** Participants completed either the nonverbal component of the Shipley-2 (Shipley et al., 2009), Cattell Culture Fair Intelligence Test (Cattell & Cattell, 1960), or the Kaufman Brief Intelligence Test (KBIT-2; Kaufman & Kaufman, 2004), as a measure of fluid intelligence. Each test was administered according to standardized instructions and converted to normed scores by a set of tables based on age with a population mean of 100 and a standard deviation of 15.

**Table 3** Means (and SDs) of background variables by age group and language group in Study 3

Age group	Younger adults		Older adults	
	Monolinguals	Bilinguals	Monolinguals	Bilinguals
<i>n</i>	26	33	26	27
Age (years)	20.5 (2.3)	20.6 (3.5)	70.9 (5.6)	71.8 (4.8)
Education (years)	12.8 (1.0)	13.2 (1.2)	15.0 (2.5)	15.9 (2.1)
Nonverbal intelligence	99.4 (11.5)	101.2 (16.3)	102.3 (13.7)	96.1 (8.7)
English vocabulary	100.9 (8.8)	94.2 (12.1)	109.2 (9.1)	107.7 (8.4)
<i>LSBQ self-ratings<sup>a</sup></i>				
Age learned English	1.2 (1.2)	6.3 (5.1)	0.9 (1.2)	7.7 (4.9)
Age learned non-English language	8.0 (3.9)	2.1 (2.8)	24.6 (26.5)	1.4 (2.8)
	<i>n</i> = 10	<i>n</i> = 33	<i>n</i> = 9	<i>n</i> = 27
English proficiency	99.9 (0.5)	90.9 (10.8)	99.3 (2.3)	96.2 (7.1)
Non-English proficiency	20.3 (11.6)	90.1 (11.7)	15.5 (11.1)	92.9 (11.5)
	<i>n</i> = 10	<i>n</i> = 33	<i>n</i> = 9	<i>n</i> = 27
English usage	99.2 (3.3)	64.4 (16.0)	99.5 (1.9)	74.1 (21.9)

<sup>a</sup> Self-report ratings of proficiency range from 0 = “no proficiency” and 100 = “fully fluent” and usage from 0 = “all other language” and 100 = “all English”. Proficiency ratings were based on an average of speaking and understanding. Usage ratings were based on an average consisting of ratings for three usage contexts. A subset of monolinguals indicated having some limited non-English language experience at some point throughout their lifetime. The non-English language profile data was not analyzed statistically across groups



**DRM task** Sixteen lists of 15 semantically associated words were selected from the 55 normed DRM lists used by Roediger, Watson, McDermott, and Gallo (2001). From these 16 lists, each participant studied eight of the lists, and the remaining eight lists were used during the recognition test as new items. To examine the effects of BAS, there were eight lists of low BAS and eight lists of high BAS words, according to the mean BAS value reported by Roediger et al. (2001). The low BAS lists ( $M = .04$ ,  $SD = .03$ , range: .01–.10) and high BAS lists ( $M = .26$ ,  $SD = .06$ , range: .20–.35) differed significantly on mean backward association strength,  $t(14) = -8.96$ ,  $p < .0001$ . A medium BAS list was used as practice. For each participant, four high BAS lists and four low BAS lists were randomly selected for study presentation. The order of presentation of words within each list was the same as Roediger et al. (2001).

The recognition test consisted of 64 items: eight critical lures from the studied lists, three items from each of the eight studied lists, and 32 new items taken from the unstudied lists. The three recognition test items that were selected from each list were based on the norms presented in Roediger et al. (2001) and indicate the strength of association of each item to the critical lure. Based on the normative data, the three items selected for the recognition test were the item most often generated as an associate to the critical lure, the sixth most often generated item in the list, and the 11th most often generated item in the list.

## Procedure

Participants began by completing the LSBQ and were then given the measure of English receptive vocabulary (Shipley vocabulary,  $n = 96$ , or PPVT,  $n = 16$ ), and the measure of nonverbal fluid intelligence (Shipley block patterns,  $n = 94$ , Cattell,  $n = 3$ , or KBIT,  $n = 13$ ). The DRM task was run on E-Prime 2.0. Items were presented visually to minimize the effects of age-related hearing loss on perception of list words.

In the first phase, participants studied eight lists (four high mean BAS and four low mean BAS), each consisting of 15 words presented serially on a computer screen, yielding a total of 120 items. Each word was displayed in black, Courier New, bold, size 18 font at the center of a blank white screen for 1,500 ms, with a fixation cross appearing between words for 1,000 ms followed by a blank white screen presented at a fixed ITI of 500 ms prior to word onset.

During the recognition phase, 64 words were presented on the computer screen, one at a time in random order, and participants pressed one of two mouse buttons on the left or right side of the monitor to indicate whether the word had been part of a study list. After a response or 5,000 ms had elapsed, the next word appeared. Right and left response keys were counterbalanced across participants.

## Results and discussion

### Background measures

Background measures are reported in Table 3. All measures were examined with two-way ANOVAs for age group and language group. Language groups were equivalent in age, but older adults had more years of education than younger adults,  $F(1, 108) = 52.35$ ,  $p < .0001$ ,  $\eta_p^2 = .33$ , and bilinguals had more years of education than monolinguals,  $F(1, 108) = 4.08$ ,  $p = .05$ ,  $\eta_p^2 = .04$ , with no significant interaction,  $F < 1$ . All groups were equivalent on nonverbal intelligence,  $F_s < 1$ . Older adults scored higher on English receptive vocabulary than younger adults,  $F(1, 108) = 33.66$ ,  $p < .0001$ ,  $\eta_p^2 = .24$ , and scores were also higher in monolinguals than in bilinguals,  $F(1, 108) = 4.78$ ,  $p = .031$ ,  $\eta_p^2 = .04$ , with no significant interaction,  $F < 1$ . Not surprisingly, monolinguals learned English at an earlier age than bilinguals did,  $F(1, 108) = 68.77$ ,  $p < .0001$ ,  $\eta_p^2 = .39$ , with no effect of group or interaction of age and group,  $F_s < 1$ . Monolinguals had higher English proficiency scores than bilinguals did,  $F(1, 108) = 21.29$ ,  $p < .0001$ ,  $\eta_p^2 = .16$ , with a larger discrepancy in younger adults,  $F(1, 57) = 17.93$ ,  $p < .0001$ ,  $\eta_p^2 = .24$ , than older adults,  $F(1, 51) = 4.63$ ,  $p = .036$ ,  $\eta_p^2 = .08$ . However, the group means for English proficiency scores were all above 90, so in the range of the population mean. There was no main effect of age for English proficiency scores,  $F < 1$ . For English usage, monolinguals used English more than bilinguals did,  $F(1, 108) = 128.84$ ,  $p < .0001$ ,  $\eta_p^2 = .54$ , and older adults used English slightly more than young adults did,  $F(1, 108) = 3.56$ ,  $p = .06$ ,  $\eta_p^2 = .03$ . The interaction between age and language group with respect to English usage was also marginally significant,  $F(1, 108) = 3.1$ ,  $p = .07$ ,  $\eta_p^2 = .03$ .

### DRM task

Recognition data are presented in Table 4. First, the percentage of ‘OLD’ responses to studied words was examined in a three-way ANOVA for age group, language group, and BAS. There were no effects of age group,  $F(1, 108) = 1.24$ ,  $ns$ , language group,  $F(1, 108) = 1.50$ ,  $ns$ , or their interaction,  $F < 1$ . There was a main effect of BAS,  $F(1, 108) = 21.15$ ,  $p < .0001$ , with more ‘OLD’ responses to high BAS words than low BAS words, but no interaction effects,  $F_s < 1$ . The accuracy of detecting the studied words was examined in a  $d$ -prime analysis comparing correct responses to studied words against false alarms (excluding critical lures). These values are also reported in Table 4. A two-way ANOVA for age group and language group indicated no significant effects or interaction, all  $F_s < 1$ .

A similar analysis examining the accuracy of responding to unstudied words was performed on ‘OLD’ responses to those words. There were no significant effects, all  $F_s < 1$ .

**Table 4** Mean percentage of ‘OLD’ responses (and standard error of the mean) for studied, unstudied, critical lure items, and *d*-prime overall and by backward association strength (BAS) by age and language group in Study 3

Group	Younger monolinguals	Younger bilinguals	Older monolinguals	Older bilinguals
Studied words	79.17 (2.77)	76.89 (2.71)	83.33 (2.29)	78.86 (3.06)
High BAS	82.05 (3.29)	79.04 (3.35)	87.82 (2.23)	83.64 (3.06)
Low BAS	76.28 (3.12)	74.75 (2.77)	78.85 (3.24)	74.07 (3.76)
Unstudied words	9.78 (2.09)	9.41 (4.16)	10.90 (4.04)	8.95 (2.65)
High BAS	9.62 (2.25)	9.22 (4.19)	10.58 (4.12)	8.95 (2.81)
Low BAS	9.94 (2.25)	9.60 (4.25)	11.22 (4.21)	8.95 (2.85)
Critical lures	74.52 (4.93)	62.88 (4.69)	81.73 (3.94)	75.00 (4.43)
High BAS	79.81 (4.60)	65.15 (6.43)	78.85 (5.13)	75.93 (5.89)
Low BAS	69.23 (6.25)	60.61 (5.11)	84.62 (4.18)	74.07 (4.11)
<i>D</i> -prime	0.69 (0.03)	0.67 (0.05)	0.72 (0.04)	0.69 (0.04)

<sup>a</sup> *D*-prime was calculated as hits (correct OLD response to studied items) minus false alarms (incorrect OLD response to new items) excluding critical lures

False alarms to critical lures were examined in an ANOVA for age group, language group, and BAS. Older adults made more false alarms than younger adults did,  $F(1, 108) = 4.45$ ,  $p = .03$ ,  $\eta_p^2 = .04$ , and monolinguals made more false alarms than bilinguals did,  $F(1, 108) = 4.02$ ,  $p = .048$ ,  $\eta_p^2 = .04$ , with no interaction,  $F < 1$ . There was no overall effect of BAS,  $F < 1$ , or significant interactions with BAS,  $F_s < 1$ . The sensitivity of the design to produce a language group difference to critical lures was evaluated using G\*Power. A one-way ANOVA for language group found a significant difference between groups,  $F(1, 110) = 4.47$ ,  $p = .03$ ,  $d = 0.41$ . To achieve power of .80, the sensitivity analysis indicated a critical  $F$  value of 3.92 and effect size of 0.21, both less than the obtained values. Therefore, the design has adequate sensitivity to detect this effect.

As noted, monolingual and bilinguals differed in English proficiency, vocabulary, and the LSBQ. None of these demographic measures were related to measures of false memory. For young adult bilinguals (monolinguals had little variability in English-language measures), correlations between these three language measures and false memory were all lower than  $r = .2$ ,  $ps > .18$ . For older adult bilinguals, correlations were all less than  $r = .3$  and  $ps > .10$ .

The results replicate those found in Experiment 2 showing more false alarms for monolinguals than for bilinguals on a DRM recognition test. The results were similar for younger and older adults, with no interaction effect. The typical age-related increases in false memory were present, but they did not interact with language status, suggesting that these factors exert independent effects on false memory. Moreover, the results extend the findings from auditory presentations used in Study 2 to visual presentations in Study 3, with no apparent difference in outcomes.

## General discussion

Across three experiments, monolinguals and bilinguals showed different susceptibility to false memories in two variants of the DRM task. In Study 1, using a phonological version of the task, bilinguals were more susceptible to critical lures than monolinguals were. However, this pattern reversed in the semantic paradigm used in Experiments 2 and 3, where bilinguals exhibited fewer false alarms than did monolinguals. Although this finding rests on a cross-experiment comparison and not a full factorial design with random assignment to conditions, the pattern for the semantic paradigm was replicated in Experiment 3. Experiment 3 also extended the results to visual presentations and included older adults. Importantly, the replication of the results in Experiments 2 and 3 were carried out independently in different sites using slightly different versions of the task and different populations, a point that enhances their generalizability.

## Attentional control mechanisms in monolingual and bilingual language processing

The primary motivation for the current study was to use the DRM paradigm as a method of interrogating language processing in bilinguals and assessing how it differs from that of monolinguals. The notion that bilingual language processing routinely requires selection of target items from jointly activated and interfering competitors has led to the suggestion that bilinguals may have an advantage relative to monolinguals in selecting competitors from among jointly activated candidates. However, the findings from Experiment 1 demonstrating increased susceptibility to phonological false memories for bilinguals compared with monolinguals suggest that, at least at the phonological level, bilingualism is not associated with improved attentional control in language selection. If

anything, the findings argue that bilinguals are at a disadvantage relative to monolinguals in the face of competing phonological activation.

What then might account for increased susceptibility to phonological false memories in bilinguals? One possibility is that bilinguals have a greater magnitude and/or increased spread of activation at the phonological level. Similarly, it may be that bilinguals demonstrate a greater reliance on phonological information than semantic information. However, we consider this explanation unlikely as most prior research with bilinguals has shown reduced lexical activation compared with monolinguals, at least as assessed with several different language tasks (Friesen et al., 2016; Gollan & Silverberg, 2001; see Bialystok, 2017, for review). Within the context of current theories of DRM false memories, one possible explanation for the increased level of phonological false memories in bilinguals is that they are more reliant on gist representations than are monolinguals. Increased levels of false memory have been attributed to overreliance on gist rather than verbatim representations in a number of other populations, including children, older adults, and individuals with Alzheimer's disease (Brainerd & Reyna, 2002; Gomes et al., 2014; Holliday et al., 2011; Obidziński & Nieznański, 2017). We emphasize here that this suggestion remains speculative until additional evidence is available to directly compare monolinguals and bilinguals on their reliance on gist versus verbatim representations. One way to test the possibility of differential reliance on the two types of memory representations would be to manipulate factors such as list length that have been shown to alter reliance on verbatim versus gist representations (Jou, Arredondo, Li, Escamilla, & Zuniga, 2017). The prediction would be that differences in phonological false memories between monolinguals and bilinguals would be reduced (or eliminated) for shorter lists lengths.

In contrast to the results of Experiment 1, Experiments 2 and 3 found that bilinguals were less susceptible than monolinguals to false memories produced by lists of semantic associates. This was the predicted outcome from IAR and was based on the idea that continually selecting from activated representations in the two languages improves bilinguals' ability to select from a set of activated lexical representations. However, the absence of an interaction between BAS and language group raises questions about this interpretation. An improved ability to select among coactivated representations would be most advantageous for high BAS lists where demands on such a selection mechanism would be greatest. Of course, it is possible that the manipulation of BAS was not sufficiently strong to produce an interaction and more extreme differences between lists of high versus low BAS would reveal such an interaction. Therefore, additional studies directly comparing lexical activation and selection in the two groups are required before we can make more definitive

conclusions regarding attentional selection in monolinguals versus bilinguals based on findings from the DRM paradigm.

One explanation for differences in semantic false memories between monolinguals and bilinguals that the current findings argue against is differences in the magnitude of initial semantic activation. Lower levels of false memory could be observed in bilinguals if initial levels of semantic activation were reduced in this population compared with monolinguals. However, if this were the case, we should have observed lower levels of correct recognition for studied items in Experiments 2 and 3. That is, if initial activation of studied items were reduced in bilinguals, they would be expected to show lower levels of correct recognition (of studied items), but monolinguals and bilinguals did not differ in correct recognition of studied items in either Experiment 2 or 3. It remains possible, however, that differences in both the magnitude and extent that activation spreads throughout semantic networks differs across the two language groups, contributing to reduced semantic false memories in bilinguals.

### Phonological versus semantic false memories

Although mechanisms mediating both semantic and phonological false memories in the DRM paradigm remain a current topic of investigation, the present findings are consistent with prior literature suggesting that the two types of false memories are generated by at least partially distinct mechanisms (Ballou & Sommers, 2008; Chan et al., 2005; Watson et al., 2003). In the current study, bilingualism was associated with an increased susceptibility to phonological false memories, but a reduced susceptibility to semantic false memories.

One potential limitation of the current study (for Experiments 2 and 3 in particular) is that sample sizes were somewhat smaller than in previous studies of both phonological and semantic false memories. However, we note that sensitivity analyses for each experiment indicated that power was sufficient to detect the key effects of interest.

In summary, the present studies demonstrated that bilingualism is associated with increased susceptibility to phonological false memories, but reduced susceptibility to semantic false memories. Although evidence for specific mechanisms leading to this pattern will need to await more direct evidence regarding mechanisms mediating the generation of false memories, the findings advance our understanding about differences between language processing in monolinguals and bilinguals. Bilinguals do not have a universal advantage on memory tasks in which selective attention is required to avoid false alarms to nonpresented items. Instead, the current findings suggest that bilinguals and monolinguals likely differ at both the phonological and semantic levels of processing. This proposal illustrates that the type of error captured by the DRM provides an ideal opportunity to advance our understanding of

language processing in bilinguals and how it differs (or does not differ) from that in monolinguals.

**Author note** Study 3 was supported by Grant A2559 from the Natural Sciences and Engineering Research Council, Canada (NSERC) to Ellen Bialystok. The remainder of this research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors wish to thank Michaela Babbitt, Niki Runge, Cari A. Bogulski, and Zehra Kamani for their assistance during data collection.

**Open practices statement** All data and materials for the three experiments are available by request.

## References

- Anderson, J., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2018). The Language and Social Background Questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, *50*(1), 250–263. doi:<https://doi.org/10.3758/s13428-017-0867-9>
- Ballou, M. R., & Sommers, M. S. (2008). Similar phenomena, different mechanisms: Semantic and phonological false memories are produced by independent mechanisms. *Memory & Cognition*, *36*(8), 1450–1459. doi:<https://doi.org/10.3758/mc.36.8.1450>
- Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., III, McDermott, K. B., & Yerys, B. E. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology*, *16*(3/5), 361–384. doi:<https://doi.org/10.1080/026432999380834>
- Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, *143*(3), 233–262. doi:<https://doi.org/10.1037/bul0000099>
- Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*(2), 290–303. doi:<https://doi.org/10.1037/0882-7974.19.2.290>
- Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, *11*(5), 164–169. doi:<https://doi.org/10.1111/1467-8721.00192>
- Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 137–148. doi:<https://doi.org/10.1037/0096-1523.25.1.137>
- Budson, A. E., Sullivan, A. L., Daffner, K. R., & Schacter, D. L. (2003). Semantic versus phonological false recognition in aging and Alzheimer's disease. *Brain and Cognition*, *51*(3), 251–261. doi:[https://doi.org/10.1016/s0278-2626\(03\)00030-7](https://doi.org/10.1016/s0278-2626(03)00030-7)
- Cattell, R. B., & Cattell, A. K. S. (1960). *Measuring intelligence with the Culture Fair tests*. Champaign, IL: Institute for Personality and Ability Testing.
- Chan, J. C., McDermott, K. B., Watson, J. M., & Gallo, D. A. (2005). The importance of material-processing interactions in inducing false memories. *Memory & Cognition*, *33*, 389–395. doi:<https://doi.org/10.3758/bf03193057>
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. doi:<https://doi.org/10.1037/0033-295X.82.6.407>
- Colomé, À. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language*, *45*(4), 721–736. doi:<https://doi.org/10.1006/jmla.2001.2793>
- Costa, A., Hernandez, M., & Sebastian-Galles, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*(1), 59–86. doi:<https://doi.org/10.1016/j.cognition.2006.12.013>
- Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, *41*(3), 365–397. doi:<https://doi.org/10.1006/jmla.1999.2651>
- De Groot, A. M. B. (1992). Bilingual lexical representation: A closer look at conceptual representations. In R. Frost & L. Katz (Eds.), *Advances in psychology, Vol. 94. Orthography, phonology, morphology, and meaning* (pp. 389–412). Oxford, England: North-Holland. doi:[https://doi.org/10.1016/s0166-4115\(08\)62805-8](https://doi.org/10.1016/s0166-4115(08)62805-8)
- Dehon, H., & Brédart, S. (2004). False memories: Young and older adults think of semantic associates at the same rate, but young adults are more successful at source monitoring. *Psychology and Aging*, *19*(1), 191–197. doi:<https://doi.org/10.1037/0882-7974.19.1.191>
- Dennis, N. A., Kim, H., & Cabeza, R. (2007). Effects of aging on true and false memory formation: An fMRI study. *Neuropsychologia*, *45*(14), 3157–3166. doi:<https://doi.org/10.1016/j.neuropsychologia.2007.07.003>
- Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*(4), 496–518. doi:<https://doi.org/10.1006/jmla.1999.2654>
- Dijkstra, T., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Scientific psychology series. Localist connectionist approaches to human cognition* (pp. 189–225). Mahwah, NJ: Erlbaum.
- Dijkstra, T., Van Heuven, W. J. B., & Grainger, J. (1998). Simulating cross-language competition with the bilingual interactive activation model. *Psychologica Belgica*, *38*(3/4), 177–196.
- Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody Picture Vocabulary Test*. Washington, DC: American Guidance Service.
- Finley, J. R., Sungkhasettee, V. W., Roediger, H. L., III, & Balota, D. A. (2017). Relative contributions of semantic and phonological associates to over-additive false recall in hybrid DRM lists. *Journal of Memory and Language*, *93*, 154–168. doi:<https://doi.org/10.1016/j.jml.2016.07.006>
- Francis, W. S. (1999). Cognitive integration of language and memory in bilinguals: Semantic representation. *Psychological Bulletin*, *125*(2), 193–222. doi:<https://doi.org/10.1037/0033-2909.125.2.193>
- Friesen, D. C., Chung-Fat-Yim, A., & Bialystok, E. (2016). Lexical selection differences in monolingual and bilingual listeners. *Brain and Language*, *152*, 1–13. doi:<https://doi.org/10.1016/j.bandl.2015.11.001>
- Gallo, D. A., & Roediger, H. L., III. (2002). Variability among word lists in eliciting memory illusions: Evidence for associative activation and monitoring. *Journal of Memory and Language*, *47*(3), 469–497. doi:[https://doi.org/10.1016/s0749-596x\(02\)00013-x](https://doi.org/10.1016/s0749-596x(02)00013-x)
- Gallo, D. A., & Roediger, H. L., III. (2003). The effects of associations and aging on illusory recollection. *Memory & Cognition*, *31*(7), 1036–1044. doi:<https://doi.org/10.3758/BF03196124>
- Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience*, *8*(10), 1298–1300. doi:<https://doi.org/10.1038/nn1543>
- Giezen, M. R., & Emmorey, K. (2017). Evidence for a bimodal bilingual disadvantage in letter fluency. *Bilingualism: Language and Cognition*, *20*(1), 42–48. doi:<https://doi.org/10.1017/s1366728916000596>
- Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging,

- bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58, 787–814. doi:<https://doi.org/10.1016/j.jml.2007.07.001>
- Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, 33(7), 1220–1234. doi:<https://doi.org/10.3758/BF03193224>
- Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish–English bilinguals. *Neuropsychology*, 16(4), 562–576.
- Gollan, T. H., Sandoval, T., & Salmon, D. P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological Science*, 22(9), 1155–1164. doi:<https://doi.org/10.1177/0956797611417002>
- Gollan, T. H., & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew-English bilinguals. *Bilingualism: Language and Cognition*, 4(1), 63–83. doi:<https://doi.org/10.1017/S136672890100013X>
- Gomes, C. F. A., Cohen, B. R., Desai, A., Brainerd, C. J., & Reyna, V. F. (2014). Aging and false memory: Fuzzy-trace theory and the elderly eyewitness. In M. P. Toglia, D. F. Ross, J. Pozzulo, & E. Pica (Eds.), *The elderly eyewitness in court* (pp. 137–166). New York, NY: Psychology Press.
- Grainger, J. (1993). Visual word recognition in bilinguals. In R. Schreuder & B. Weltens (Eds.), *Studies in bilingualism, Vol. 6: The bilingual lexicon* (pp. 11–25). Amsterdam, Netherlands: John Benjamins. doi:<https://doi.org/10.1075/sibil.6.03gra>
- Green, D. W. (1993). Towards a model of L2 comprehension and production. In R. Schreuder & B. Weltens (Eds.), *Studies in bilingualism, Vol. 6: The bilingual lexicon* (pp. 249–277). Amsterdam, Netherlands: John Benjamins. doi:<https://doi.org/10.1075/sibil.6.12gre>
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press. doi:[https://doi.org/10.1016/s0079-7421\(08\)60041-9](https://doi.org/10.1016/s0079-7421(08)60041-9)
- Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, 1(3), 213–229. doi:<https://doi.org/10.1017/s1366728998000364>
- Holliday, R. E., Brainerd, C. J., & Reyna, V. F. (2011). Developmental reversals in false memory: Now you see them, now you don't! *Developmental Psychology*, 47(2), 442–449. doi:<https://doi.org/10.1037/a0021058>
- Holliday, R. E., & Weekes, B. S. (2006). Dissociated developmental trajectories for semantic and phonological false memories. *Memory*, 14(5), 624–636. doi:<https://doi.org/10.1080/09658210600736525>
- Howe, M. L., Wimmer, M. C., & Blease, K. (2009). The role of associative strength in children's false memory illusions. *Memory*, 17(1), 8–16. doi:<https://doi.org/10.1080/09658210802438474>
- Jou, J., Arredondo, M. L., Li, C., Escamilla, E. E., & Zuniga, R. (2017). The effects of increasing semantic-associate list length on the Deese–Roediger–McDermott false recognition memory: Dual false-memory process in retrieval from sub- and supraspan lists. *Quarterly Journal of Experimental Psychology (Hove)*, 70(10), 2076–2093. doi:<https://doi.org/10.1080/17470218.2016.1222446>
- Ju, M., & Luce, P. A. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science*, 15(5), 314–318. doi:<https://doi.org/10.1111/j.0956-7976.2004.00675.x>
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* (2nd ed.). Bloomington, MN: Pearson.
- Kaushanskaya, M., & Marian, V. (2007). Bilingual language processing and interference in bilinguals: Evidence from eye tracking and picture naming. *Language Learning*, 57(1), 119–163. doi:<https://doi.org/10.1111/j.1467-9922.2007.00401.x>
- Kroll, J. F. (1993). Accessing conceptual representations for words in a second language. In R. Schreuder & B. Weltens (Eds.), *Studies in bilingualism, Vol. 6: The bilingual lexicon* (pp. 53–81). Amsterdam, Netherlands: John Benjamins. doi:<https://doi.org/10.1075/sibil.6.05kro>
- Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, 9(2), 119–135. doi:<https://doi.org/10.1017/S1366728906002483>
- Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. *Annual Review of Linguistics*, 1(1), 377–394. doi:<https://doi.org/10.1146/annurev-linguist-030514-124937>
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149–174. doi:<https://doi.org/10.1006/jmla.1994.1008>
- Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science*, 20(12), 1507–1515. doi:<https://doi.org/10.1111/j.1467-9280.2009.02480.x>
- Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, 114(1), 29–41. doi:<https://doi.org/10.1016/j.cognition.2009.08.014>
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, 19(1), 1–36. doi:<https://doi.org/10.1097/00003446-199802000-00001>
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940–967. doi:[https://doi.org/10.1044/1092-4388\(2007\)067](https://doi.org/10.1044/1092-4388(2007)067)
- Marian, V., & Spivey, M. (2003a). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, 24(2), 173–193. doi:<https://doi.org/10.1017/s0142716403000092>
- Marian, V., & Spivey, M. (2003b). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, 6(2), 97–115. doi:<https://doi.org/10.1017/s1366728903001068>
- McDowd, J. M., & Filion, D. L. (1992). Aging, selective attention, and inhibitory processes: A psychophysiological approach. *Psychology and Aging*, 7(1), 65–71. doi:<https://doi.org/10.1037/0882-7974.7.1.65>
- McGeown, S. P., Gray, E. A., Robinson, J. L., & Dewhurst, S. A. (2014). What factors underlie children's susceptibility to semantic and phonological false memories? Investigating the roles of language skills and auditory short-term memory. *Cognition*, 131(3), 323–329. doi:<https://doi.org/10.1016/j.cognition.2014.02.005>
- Morford, J. P., Wilkinson, E., Villwock, A., Piñar, P., & Kroll, J. F. (2011). When deaf signers read English: Do written words activate their sign translations? *Cognition*, 118(2), 286–292. doi:<https://doi.org/10.1016/j.cognition.2010.11.006>
- Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford, UK: Oxford University Press.
- Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition*, 25(6), 838–848. doi:<https://doi.org/10.3758/bf03211328>
- Norris, D., Cutler, A., McQueen, J. M., & Butterfield, S. (2006). Phonological and conceptual activation in speech comprehension. *Cognitive Psychology*, 53(2), 146–193. doi:<https://doi.org/10.1016/j.cogpsych.2006.03.001>

- Obidziński, M., & Nieznański, M. (2017). False memory for orthographically versus semantically similar words in adolescents with dyslexia: A fuzzy-trace theory perspective. *Annals of Dyslexia*, 67(3), 318–332. doi:<https://doi.org/10.1007/s11881-017-0146-6>
- Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232–258. doi:<https://doi.org/10.1016/j.cogpsych.2012.12.002>
- Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory and Language*, 35(2), 261–285. doi:<https://doi.org/10.1006/jmla.1996.0015>
- Potter, M. C., So, K. F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23(1), 23–38. doi:[https://doi.org/10.1016/s0022-5371\(84\)90489-4](https://doi.org/10.1016/s0022-5371(84)90489-4)
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 803–814. doi:<https://doi.org/10.1037/e537272012-273>
- Roediger, H. L., III, McDermott, K. B., Pisoni, D. B., & Gallo, D. A. (2004). Illusory recollection of voices. *Memory*, 12(5), 586–602. doi:<https://doi.org/10.1080/09658210344000125>
- Roediger, H. L., III, McDermott, K. B., & Robinson, K. J. (1998). The role of associative processes in creating false memories. In M. A. Conway, S. E. Gathercole, & C. Cornoldi (Eds.), *Theories of memory II* (pp. 187–245). Hove, England: Psychological Press.
- Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, 8(3), 385–407. doi:<https://doi.org/10.3758/bf03196177>
- Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, 13(2), 231–252. doi:<https://doi.org/10.1017/s1366728909990514>
- Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology*, 49(1), 289–318. doi:<https://doi.org/10.1146/annurev.psych.49.1.289>
- Schulpen, B., Dijkstra, T., Schriefers, H. J., & Hasper, M. (2003). Recognition of interlingual homophones in bilingual auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 29(6), 1155–1178. doi:<https://doi.org/10.1037/0096-1523.29.6.1155>
- Shiple, W. C., Gruber, C. P., Martin, T. A., & Klein, A. M. (2009). *Shiple-2 manual*. Los Angeles, CA: Western Psychological Services.
- Snodgrass, J. G. (1984). Concepts and their surface representations. *Journal of Memory and Language*, 23(1), 3–22. doi:[https://doi.org/10.1016/s0022-5371\(84\)90479-1](https://doi.org/10.1016/s0022-5371(84)90479-1)
- Sommers, M. S., & Huff, L. M. (2003). The effects of age and dementia of the Alzheimer's type on phonological false memories. *Psychology and Aging*, 18(4), 791–806. doi:<https://doi.org/10.1037/0882-7974.18.4.791>
- Sommers, M. S., & Lewis, B. P. (1999). Who really lives next door: Creating false memories with phonological neighbors. *Journal of Memory and Language*, 40(1), 83–108. doi:<https://doi.org/10.1006/jmla.1998.2614>
- Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, 10(3), 281–284. doi:<https://doi.org/10.1111/1467-9280.00151>
- Sullivan, M. D., Poarch, G. J., & Bialystok, E. (2018). Why is lexical retrieval slower for bilinguals? Evidence from picture naming. *Bilingualism: Language and Cognition*, 21(3), 479–488. doi:<https://doi.org/10.1017/S1366728917000694>
- Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, 104(30), 12530–12535. doi:<https://doi.org/10.1073/pnas.0609927104>
- Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging*, 13(2), 230–241. doi:<https://doi.org/10.1037/0882-7974.13.2.230>
- Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: A review of meta-analyses. *Neuroscience & Biobehavioral Reviews*, 26(7), 849–857. doi:[https://doi.org/10.1016/s0149-7634\(02\)00071-4](https://doi.org/10.1016/s0149-7634(02)00071-4)
- von Bastian, C. C., Souza, A. S., & Gade, M. (2016). No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General*, 145(2), 246–258. doi:<https://doi.org/10.1037/xge0000120>
- Wallace, W. P., Stewart, M. T., & Malone, C. P. (1995). Recognition memory errors produced by implicit activation of word candidates during the processing of spoken words. *Journal of Memory & Language*, 34(4), 417–439.
- Watson, J. M., Balota, D. A., & Roediger, H. L., III. (2003). Creating false memories with hybrid lists of semantic and phonological associates: Over-additive false memories produced by converging associative networks. *Journal of Memory and Language*, 49(1), 95–118. doi:[https://doi.org/10.1016/s0749-596x\(03\)00019-6](https://doi.org/10.1016/s0749-596x(03)00019-6)
- Watson, J. M., Balota, D. A., & Sergent-Marshall, S. D. (2001). Semantic, phonological, and hybrid veridical and false memories in healthy older adults and in individuals with dementia of the Alzheimer type. *Neuropsychology*, 15(2), 254–267. doi:<https://doi.org/10.1037/0894-4105.15.2.254>
- Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, 50(1), 1–25. doi:[https://doi.org/10.1016/s0749-596x\(03\)00105-0](https://doi.org/10.1016/s0749-596x(03)00105-0)
- Westbury, C., Buchanan, L., & Brown, N. (1999). Sounds of the neighbourhood: False memories and the structure of the phonological lexicon. *Journal of Memory and Language*, 46(3), 622–651. doi:<https://doi.org/10.1006/jmla.2001.2821>
- Zhao, X., & Li, P. (2010). Bilingual lexical interactions in an unsupervised neural network model. *International Journal of Bilingual Education and Bilingualism*, 13(5), 505–524. doi:<https://doi.org/10.1080/13670050.2010.488284>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.