Bilingualism: Language and Cognition

cambridge.org/bil

Performance difference in verbal fluency in bilingual and monolingual speakers

Abhijeet Patra[1], Arpita Bose[1] and Theodoros Marinis[1,2]

[1]School of Psychology and Clinical Language Sciences, University of Reading, UK and [2]Department of Linguistics, University of Konstanz, Germany

Research Article

Cite this article: Patra A, Bose A, Marinis T (2020). Performance difference in verbal fluency in bilingual and monolingual speakers. *Bilingualism: Language and Cognition* **23**, 204–218. https://doi.org/10.1017/S1366728918001098

Received: 15 March 2018
Revised: 30 October 2018
Accepted: 5 November 2018
First published online: 19 February 2019

Key words:
semantic fluency; letter fluency; executive control; inhibition; time-course analysis; Bengali

Address for correspondence: Arpita Bose, Ph.D., Email: a.bose@reading.ac.uk

Abstract

Research has shown that bilinguals can perform similarly, better or poorly on verbal fluency task compared to monolinguals. Verbal fluency data for semantic (animals, fruits and vegetables, and clothing) and letter fluency (F, A, S) were collected from 25 Bengali–English bilinguals and 25 English monolinguals in English. The groups were matched for receptive vocabulary, age, education and non-verbal intelligence. We used a wide range of measures to characterize fluency performance: number of correct, fluency difference score, time-course analysis (1st RT, Sub-RT, initiation, slope), clustering, and switching. Participants completed three executive control measures tapping into inhibitory control, mental-set shifting and working memory. Differences between the groups were significant when executive control demands were higher such as number of correct responses in letter fluency, fluency difference score, Sub-RT, slope and cluster size for letter fluency, such that bilinguals outperform the monolinguals. Stroop performance correlated positively with the slope only for the bilinguals.

Introduction

The literature is abuzz with arguments for and against the linguistic and executive control differences between monolingual and bilingual speakers (Gollan, Montoya, Fennema-Notestine & Morris, 2005; Ivanova & Costa, 2008; Luo, Luk & Bialystok, 2010; Paap, Myuz, Anders, Bockelman, Mikulinsky & Sawi, 2017; Prior & MacWhinney, 2010). Previous studies have shown bilingual disadvantages in various linguistic tasks such as picture naming (Gollan et al., 2005), verbal fluency (Rosselli, Ardila, Araujo, Weekes, Caracciolo, Padilla & Ostrosky-Solí, 2000), word identification through noise (Rogers, Lister, Febo, Besing & Abrams, 2006). According to the weaker link hypothesis, the reason for bilingual disadvantage in the linguistic domain is the lesser usage of each language of a bilingual speaker resulting in weaker links between the two languages (Michael & Gollan, 2005). A sensorimotor account (Hernandez & Li, 2007) attributes the bilingual disadvantage to the delay in age of acquisition of the second language. Further, bilinguals face greater lexical competition compared to monolinguals as both languages are active during language processing (Costa & Caramazza, 1999) and the poorer performance in the linguistic domain can be attributed to this increased lexical competition (Inhibitory control model, Green, 1998).

In contrast to the disadvantages of being bilingual in the verbal domain, effect of bilingualism on executive control mechanism is hotly debated. Researchers have shown advantages (Bialystok, Craik, Klein & Viswanathan, 2004; Prior & MacWhinney, 2010) as well as no differences across various executive control tasks (Kousaie & Phillips, 2012; Paap & Greenberg, 2013; Paap et al., 2017). For example, studies have reported bilingual advantage on inhibitory control tasks (e.g., Simon task in Bialystok et al., 2004; Flanker task in Emmorey, Luk, Pyers & Bialystok, 2008), no difference or similar performance between bilinguals and monolinguals has also been noted (e.g., Stroop in Kousaie & Phillips, 2012; Flanker task and Simon task in Paap & Greenberg, 2013; Paap & Sawi, 2014). Similarly, on mental set shifting measure using colour-shape task switching paradigm have reported divergent findings ranging from advantage for bilinguals (Prior & Gollan, 2011; Prior & MacWhinney, 2010) to no differences between the two groups (Paap & Greenberg, 2013; Paap & Sawi, 2014; Paap et al., 2017). As could be seen from the literature, it is still unresolved whether bilinguals would show specific advantages on certain domains of executive control as the difference between the groups depends on cultural differences, small sample size, inappropriate statistical analysis, and the tasks used (Paap, Johnson & Sawi, 2014).

A prevalent approach in the literature has been to use separate measures of language production and executive control mechanisms; linguistic tasks tapping into language production, and non-linguistic tasks tapping into executive control processes. A better approach to inform the debate on – disadvantage or advantage – amongst language and executive control processes for bilingual and monolingual speakers would be to use a task (e.g., verbal fluency task) that simultaneously draw upon both these processes. With the exception of a handful of studies, the

© Cambridge University Press 2019. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

CAMBRIDGE UNIVERSITY PRESS

role of executive control during language production amongst bilinguals and monolinguals has not been explored (e.g., Bialystok, Craik & Luk, 2008; Friesen, Luo, Luk & Bialystok, 2015).

Researchers have used the verbal fluency task – the ability to produce as many unique words as possible in a fixed amount of time, according to a given criterion (e.g., semantic or category; letter or phonemic) – to inform the debate of linguistic and executive control differences between monolingual and bilingual speakers (Luo et al., 2010; Paap et al., 2017; Sandoval, Gollan, Ferreira & Salmon, 2010). Performance in the semantic fluency condition resembles to our day to day language activities: for example, in a semantic fluency task, participants are asked to generate items belonging to the category of clothing, and participants try to remember the items from their wardrobes. Therefore, participants can revisit the existing links in their mental lexicon related to a concept while generating novel words in the semantic fluency condition (Friesen et al., 2015). However, letter fluency condition becomes more challenging as it requires producing words starting with a letter or phoneme, which is not commonly practiced in our everyday life. Successful performance in the letter fluency condition requires coming up with strategies and suppression of the activation of related semantic concepts (e.g., Friesen et al., 2015; Luo et al., 2010). Thus, the respective contributions of linguistic and executive components are differential for semantic and letter fluency conditions: higher demands are placed on executive control mechanisms in letter fluency, while a greater emphasis is placed on linguistic abilities in semantic fluency (Delis, Kaplan & Kramer, 2001; Luo et al., 2010; Paap et al., 2017; Sandoval et al., 2010; Shao, Janse, Visser & Meyer, 2014).

Verbal fluency research comparing bilingual and monolingual performance have shown mixed results (Bialystok et al., 2008; Luo et al., 2010; Paap et al., 2017; Sandoval et al., 2010). In semantic fluency, monolinguals generate a larger number of correct responses than bilinguals (Gollan, Montoya & Werner, 2002; Rosselli et al., 2000; Sandoval et al., 2010). However, this bilingual disadvantage disappears when the groups are matched on receptive vocabulary (Bialystok et al., 2008; Luo et al., 2010). For letter fluency, findings have been wide ranging from fewer to equivalent to greater number of correct responses by bilinguals (Bialystok et al., 2008; Kormi-Nouri, Moradi, Moradi, Akbari-Zardkhaneh & Zahedian, 2012; Rosselli et al., 2000; Luo et al., 2010; Paap et al., 2017; Sandoval et al., 2010). Luo et al. (2010) found that vocabulary matched bilinguals outperform monolinguals on letter fluency, proposing that it is suggestive of better executive control in bilinguals. However, Paap et al. (2017) were unable to replicate these results. They strongly argued that "relatively better performance by a group on letter fluency compared to category fluency cannot be taken as evidence that the group has superior executive functions. Rather such a claim must be backed up by an independent and direct test of EF ability" (Paap et al., 2017, p.108). Importantly, studies exploring the relationship of independent measures of executive control and verbal fluency performance (at least in monolinguals) did not find a stronger relationship between executive control measures and the performance in letter fluency compared to semantic fluency task (Shao et al., 2014). With a limited number of empirical studies and difficulties with replication, it remains an open question whether bilinguals and monolinguals: (1) perform differently in semantic and letter fluency tasks; (2) whether their performance differences would be mediated by specific aspects of executive control abilities.

Moving beyond the number of correct responses, we used a wide range of variables to characterize verbal fluency performance, such as time-course, clustering, and switching analyses for both semantic and letter fluency (Luo et al., 2010; Troyer, Moscovitch & Winocur, 1997). Table 1 provides description of the variables and the components of verbal fluency they are assumed to index. To our knowledge, this is the first study that systemically compares healthy bilinguals and monolinguals on this full range of measures. In addition, we included independent measures of executive processes (i.e., inhibition, shifting and memory) to compare performance differences between bilinguals and monolinguals and their relationship to verbal fluency performance. This allows us to establish if bilinguals will evidence exaggerated differences on the verbal fluency parameters that depend more on the executive component of the task and if bilinguals' better performance in letter fluency found in some studies can be attributed to differences in executive control.

As the verbal fluency task places a premium on rapid search and retrieval, temporal measures of performance, such as time-course analysis (i.e., production time of each word as a function of its position in the sequence), provide insights into the linguistic and executive control strategies (e.g., Crowe, 1998; Luo et al., 2010; Sandoval et al., 2010). In time-course analysis, the number of words generated over the 60 second time interval is grouped into 5-second time bins, with declining response rate presented by plotting the number of words produced as a function of time. Four parameters are generated from this graph: First-Response Time (1st-RT), Subsequent-Response Time (Sub-RT); initiation parameter; and slope (see Table 1 for the definition of these measures). Luo et al. (2010) compared semantic and letter fluency performance for a group of young monolinguals and two groups of young bilinguals (high-vocabulary bilinguals who were matched with monolinguals; low-vocabulary bilinguals). In letter fluency, the high-vocabulary bilinguals produced a profile of larger number of correct responses, a longer Sub-RT, and a flatter slope than the monolinguals. Similar results have been obtained by Friesen et al. (2015), who found no difference between bilinguals and monolinguals on the semantic fluency condition, but a greater number of correct responses on the letter fluency by the bilinguals.

In contrast, studies have shown that bilinguals produced longer Sub-RT along with fewer number of correct responses compared to monolinguals in letter fluency (Sandoval et al., 2010). These authors argued that the bilingual disadvantage results from cross-linguistic interference which slows down their word retrieval process, as denoted by longer Sub-RT. It has been argued that as vocabulary-matched bilinguals produced a greater number of correct responses compared to monolinguals, it is unlikely that the retrieval-slowing hypothesis can explain the bilingual advantage (Friesen et al., 2015; Luo et al., 2010). Instead, they suggest that bilinguals' better performance in the letter fluency in conjunction with the longer Sub-RT is a result of bilinguals' superior executive control abilities, which is proposed to be a by-product of constant cross-linguistic interference faced by bilinguals (Abutalebi & Green, 2008; Friesen et al., 2015; Luo et al., 2010).

The Fluency Difference Score (FDS) has been suggested to further capture the role of executive control in fluency task (Friesen et al., 2015). The FDS is calculated as the difference in the number of correct responses between the semantic and letter fluency conditions as a proportion of correct responses in the semantic fluency condition. Therefore, individuals who can maintain better

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

Table 1. Description of the verbal fluency variables and their relative contribution to the linguistic and executive control components

Parameters	Description	Linguistic component	Executive control component
Quantitative analysis			
Number of correct responses (CR)	Number of responses produced in one minute excluding any errors. Measures word retrieval abilities.	√	√
Fluency Difference Score (FDS)[1]	Measures the ability to maintain the performance in the demanding condition (i.e., letter fluency).		√
Time-course analysis[2,3]			
1st RT	Preparation time to initiate the first response.	√	
Sub-RT	Estimate for mean retrieval latency and represents the time point at which half of the total responses have been generated.		√
Initiation parameter	Measures the initial linguistic resources or vocabulary available to perform the task.	√	
Slope	Reflects how resources are monitored and used over time during the retrieval process; largely determined by executive processes.		√
Qualitative analysis[4]			
Cluster size	Strategic process that helps generating words within a subcategory and utilizes the speaker's ability to access words within subcategories.	√	
Number of switches	Strategic process to shift efficiently to a new subcategory when a subcategory is exhausted.		√

[1]Friesen et al., 2015; [2]Luo et al., 2010; [3]Sandoval et al., 2010; [4]Troyer et al., 1997.

performance in the difficult letter fluency condition would show a smaller FDS score; this is indicative of better executive control abilities (Friesen et al., 2015).

The production of words during verbal fluency performance is not evenly distributed over time but tends to be produced in "spurts" or temporal clusters, with a short time interval between words in a cluster and a longer pause between clusters (Gruenewald & Lockhead, 1980; Troyer et al., 1997). On semantic fluency tasks, the words that comprise these temporal clusters tend to be semantically related (e.g., first name farm animals, then switch to pets, then to birds); on letter fluency tasks, the words tend to be phonologically related (e.g., words that start with same first two letters, then switch to words that rhyme, then to words that have the same ending). This response pattern has led to the suggestion that performance involves two processes: a search for subcategories which corresponds to a pause between clusters followed by an output mechanism to produce as many words as possible from the subcategories (Gruenewald & Lockhead, 1980; Tröster, Fields, Testa, Paul, Blanco, Hames, Salmon & Beatty, 1998). The metrics of switching and clustering have been suggested to quantify the above two processes (Troyer et al., 1997). Specifically, clustering involves accessing and using the word store and cluster size is a measure of the ability to access words within the subcategory. Switching involves search processes and is a measure of the ability to shift efficiently from one subcategory to another; reduced switching has been attributed to executive function difficulty to shift between subcategories (Troyer, Moscovitch, Winocur, Alexander & Stuss, 1998). Both clustering and switching abilities contribute to the total number of correct responses; however, in category fluency, clustering accounts for more of the variance for number of correct responses, whilst in letter fluency, switching accounts for more of the variance for number of correct responses (Troyer et al., 1997). Thus, clustering and switching analyses provide another well-established mean to further inform the linguistic and executive debate for bilinguals vs. monolinguals.

To the best of our knowledge, no research has reported the relationship of independent executive control measures to bilingual vs. monolingual performance difference on verbal fluency. Only one study with healthy monolingual adults investigated the 60-seconds verbal fluency performance with measures of executive control (Shao et al., 2014). Shao et al. had assessed older Dutch speakers on both semantic and letter fluency conditions and related their performance with the measures of executive control (i.e., updating of working memory, operation span; inhibitory control, stop-signal task). Results revealed that only working memory ability predicted the number of correct responses in both fluency conditions. Shao et al. noted that "there was no evidence that executive control had a stronger effect on performance in the letter than in the category fluency task" (Shao et al., 2014, p. 8). The authors cautioned that the inhibitory control task (i.e., stop-signal task) used in their study may not have represented the inhibitory control required for the verbal fluency task. The stop-signal task measures how fast an individual can stop a planned response, whereas, in verbal fluency, participants need to suppress the activation of competitor lexical items (selective inhibition) to produce the target word.

For the present study, we adopted the framework developed by Miyake and his colleagues (Miyake, Friedman, Emerson, Witzki, Howerter & Wager, 2000; Miyake & Friedman, 2012) to measure the three executive control components. This framework proposes that the three executive control components share a common executive functioning factor, which is the ability to actively maintain task-related goals while controlling the lower level processing using the task-related information (Miyake & Friedman, 2012). Specifically, this is what we measured: inhibitory control (ability to inhibit the automatic, dominant, or prepotent responses when required), mental set-shifting (ability to shift between different tasks, rules, or mental representations), and working memory (constant updating and manipulation of relevant incoming information while replacing old irrelevant information). We

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

used the Stroop task to measure selective inhibition (Scott & Wilshire, 2010), the colour-shape switch task to measure the mental-set shifting ability (Prior & MacWhinney, 2010), and the backward digit span test to measure working memory (Wechsler, 1997).

Research in bilingualism has identified various factors, such as language combination of bilinguals and their language proficiency, which can confound the results. Studies including bilinguals with a range of different language combinations lead the individual variability and can result in a wider range of performance that could be attributed to typological, structural, and cultural differences amongst the languages (Eng, Vonk, Salzberger & Yoo, 2018; Marian, 2008). Inclusion of bilinguals with the same language combination allows controlling for within-group performance variation due to differences in the second language they speak. Language proficiency of bilinguals has also been shown to be also a significant contributor for verbal fluency performance (Bialystok et al., 2008; Gollan et al., 2002; Luo et al., 2010). When bilinguals are matched with monolinguals in terms of language proficiency, they either outperform (Luo et al., 2010) or perform at par with the monolinguals (Bialystok et al., 2008; Paap et al., 2017). In contrast, low proficient bilinguals perform poorly (Gollan et al., 2002) compared to the monolinguals. Therefore, it is crucial to match the bilinguals to the monolinguals in terms of language proficiency. In the present research, we have included a homogenous group of bilinguals in terms of language combination and proficiency, which we hope would decrease the within-group variability and findings could be attributed to the processes that are tested.

The current study

We compared the difference in verbal fluency performance in two groups of young healthy participants: 25 Bengali–English bilinguals and 25 English monolinguals. The groups were matched on receptive vocabulary, years of education, and non-verbal intelligence. We collected semantic (animals, fruits, vegetables) and letter (F, A, S) fluency data for 60 seconds in English. We provided detailed characterization of our bilingual participants on relevant variables for bilingualism: language history and acquisition patterns, usage patterns, proficiency, dominance, and switching habits. Our bilingual participants formed a relatively homogenous group of balanced bilinguals in terms of language of instruction during education, self-rated language proficiency, and language dominance. All bilingual participants were born in the Bengali speaking region in India and acquired Bengali as their first language. However, they currently lived in the UK and they used English more frequently than Bengali in their everyday life.

We quantified the verbal fluency performance in terms of quantitative (number of correct responses; FDS); time-course (1st-RT; Sub-RT; initiation parameter; slope); and qualitative (cluster size; number of switches). Executive control processes were measured using the Stroop (measured selective inhibition), the colour-shape switch task (measured shifting between mental sets), and the backward digital span (measured working memory) tasks.

We formulated our hypothesis from the theoretical accounts (weaker link hypothesis, inhibitory control model) described earlier. Bilingual participants in the present study were matched in vocabulary with the monolingual group. Further, our bilingual participants used English in their day-to-day life more often than Bengali. We predicted that controlling for these factors (vocabulary and usage), bilinguals would be able to perform at par with the monolinguals if bilinguals can resolve their increased cross-linguistic competition. Moreover, they might be able to perform better in linguistic conditions that require higher executive control processing (e.g., letter fluency condition). The research aims, and predictions, were as follows:

1. To determine differences in verbal fluency performance (quantitative, time course, and qualitative analysis) between bilingual and monolingual participants.

   As the groups were matched on vocabulary, we predicted bilinguals would perform similarly to monolinguals on the semantic fluency condition, but potentially produce a larger number of words than monolinguals in the letter fluency condition. In similar vein, we did not expect differences in cluster size. If bilinguals were to show superior executive control, we would expect bilinguals to demonstrate smaller FDS, more number of switches and longer Sub-RT, and flatter slope in letter fluency compared to monolinguals.
2. To determine measures of executive control (inhibitory control, mental set shifting, and working memory) that mediate verbal fluency performance difference between the groups. We expected that if bilinguals were to show an advantage in the letter fluency condition, then executive control measures would have a stronger correlation with performance measures that relate to the executive control abilities (i.e., FDS, slope, number of switches).

Methods

Participants

Twenty-five Bengali–English bilingual healthy adults ($M = 32.84$, $SD = 4.78$) and 25 English monolingual healthy adults ($M = 30.4$, $SD = 8.2$) participated in this study. Participants reported themselves to be right-handed, with normal or corrected vision, no history of hearing impairment, and no history of any neurological illness.

All participants were residing in the Berkshire county of the United Kingdom. Demographic details (age, gender, and years of education) and scores on nonverbal IQ from the Raven's standard progressive matrices *plus* version (SPM *Plus*, Raven, 2008) are presented in Table 2. Participants were also assessed on two standardised tests of receptive vocabulary: the Oxford Placement Test (Oxford University Press and Cambridge ESOL, 2001) and the British Picture Vocabulary Scale III (BPVS-III; Dunn, 2009). The groups did not differ on age, gender distribution (bilinguals: 11 females and 14 males; monolinguals: 12 males, 13 females; $p = .78$), years of education, non-verbal IQ and receptive vocabulary (see Table 2). Bilingual participants were recruited from the local Bengali community (e.g., Bengali Cultural Society of Reading). Bilinguals were immigrants who have lived in the UK, ranging from 1 year to 15 years ($M = 7.48$, $SD = 3.58$). They spoke Bengali and English fluently, had minimal or no knowledge of any other language. Monolingual participants were recruited from the university student population, who received course credit for participation and local community. Monolingual participants used only English in their day-to-day life and were functionally fluent only in English. Participants provided written consent and their participation was voluntary. The University of

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

Table 2. Mean (M), standard deviations (SD), and statistical results of the demographic variables, Raven's SPM-plus and vocabulary tests

	Bilingual ($N = 25$)		Monolingual ($N = 25$)		
Measures	*M*	*SD*	*M*	*SD*	Statistical results
Age (Years)	32.8	4.8	30.4	8.2	$t(48) = 1.3$, $p = .21$
Years of education	18.1	1.6	17.1	1.2	$U^2 = 311.5$, $p = .98$
Raven's IQ[1]	43.5	3.8	43	5.4	$U^2 = 275.5$, $p = .47$
OPT[3]	53.1	3.4	54.1	3.4	$U^2 = 251.5$, $p = .23$
BPVS III[4]	157.8	4.8	159.8	4.6	$U^2 = 269.5$, $p = .40$

[1] – Raven's Standard Progressive Matrices Intelligence Quotients (Raven, 2008), maximum score possible 60, greater score indicates higher non-verbal intelligence; [2] – Mann-Whitney U test; [3] – Oxford Quick Placement Test (2001), maximum possible score was 60, higher score indicates higher receptive vocabulary; [4] – British Picture Vocabulary Scale, Third Edition (Dunn & Dunn, 2009), maximum possible score was 164, higher score indicates higher receptive vocabulary.

Reading Research Ethics Committee approved all the experimental procedures.

Measures of bilingualism

Bilinguals were assessed using various measures to characterize their bilingualism. We adapted and modified the questionnaire developed by Muñoz, Marquardt & Copeland (1999). This questionnaire assessed language acquisition history, instruction of language during education, self-rated language proficiency (in speaking, comprehension, reading and writing), and the current language usage pattern. Language dominance was measured using the language dominance questionnaire (Dunn & Tree, 2009) and language switching habits were assessed using a language switching questionnaire (Rodriguez-Fornells, Krämer, Lorenzo-Seva, Festman & Münte, 2012). All the questionnaires are provided as Supplementary Material (Appendix S1, Supplementary Materials).

There was no significant difference amongst bilinguals' Bengali and English on the language of instruction during education, subjective language proficiency ratings (speaking, comprehension, reading, and writing abilities) and language dominance. This indicated a balanced bilingualism on these domains. However, during childhood, bilinguals had significantly greater Bengali exposure during acquisition ($M = 14.3$, $SD = 2.6$) than English ($M = 2.5$, $SD = 2.3$). Current usage of language was predominantly English; they were more prone to switch from Bengali to English than the reverse during day-to-day communication.

Verbal fluency measures

Trials and procedures

Participants completed two verbal fluency conditions – semantic and letter – in English. They were asked to produce as many words as possible in 60 seconds. In the semantic condition, participants produced words in three categories – animal, fruits and vegetables, and clothing items. In the letter condition, participants were asked to produce words that start with letters *F, A,* and *S*. The restrictions for the letter conditions were to produce unique words that are not proper names or not numbers (e.g., *Singapore, seven*), and to not produce variants of the same words (e.g., *shop, shopper, shopping*). The order of the fluency conditions was randomized across participants; however, the trials were blocked by condition. Each participant was tested individually in a quiet room. After providing the instruction, the participant started a trial only when the tester said "start". This ensured that there was a definitive starting point for each trial. Responses were recorded with a digital voice recorder and later analysed for the following variables.

Data coding and analysis

All responses (including repetition and errors) were transcribed verbatim. Each correct response was time-stamped using PRAAT (Boersma & Weenink, 2015). The time-stamping enabled us to index the onset of a response from the onset of the trial (i.e., "start"), which allowed us to calculate the variables in time-course analysis. We measured the following variables for each trial:

1. Number of correct responses (CR): the number of responses produced in one-minute excluding errors. In semantic condition, errors were repetition of same words, words that were not from the target category (e.g., *cat* as a response for clothing category), and cross-linguistic intrusions. In letter condition, errors were repetition of same words, words that began with a different letter (e.g., *pig* as a response for letter *F*), proper names (e.g., *France* as a response to letter *F*), same word but with inflectional or derivational suffixes (e.g., *fast*, *faster*, *fastest* were counted as single CR), and cross-linguistic intrusions.
2. Fluency Difference Score (FDS): the differences in the number of correct responses between semantic and letter fluency conditions as a proportion of correct responses in the semantic fluency condition.

$$FDS = \left(CR\ semantic\ fluency - CR\ letter\ fluency\right) / CR\ semantic\ fluency$$

3. Time-course analysis: four variables – 1st RT; Sub-RT; initiation parameter; and slope – were computed based on the timing of the responses (Luo et al., 2010). Based on the time tag, CRs were grouped into 5 sec bins over each 60 sec trial, resulting in 12 bins. The group means of CR in each of the twelve bins were calculated for each semantic and letter fluency trial. The means of CRs for each trial were plotted using a line graph (x variable, bins; y-variable, mean CR). This graph was then fitted with a logarithmic function. An example of a logarithmic function is $y = 4.39–1.41\ In(t)$, where y is the estimated value of the function at different points of time(t). Two central measures derived from this plot were: initiation parameter and slope.

First-RT (1st-RT) is the time interval from the beginning of the trial to the onset of first response. The first response usually takes longer than the subsequent responses and this delay in

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

first response has been linked to the task preparation (Rohrer, Wixted, Salmon & Butters, 1995).

Subsequent-RT (Sub-RT) is the average value of the time intervals from the onset of first response to the onset of each subsequent response. Thus, Sub-RT provides a good estimate for mean retrieval latency and represents the time point at which half of the total responses have been generated (Sandoval et al., 2010). A longer mean Sub-RT indicates that performance extends later into the time course, but interpretation of this variable depends on the total number of correct responses (Luo et al., 2010). If one group produces more correct responses than another group and has longer mean Sub-RT, then the interpretation is that the group has superior control (and equivalent or better vocabulary) and could continue generating responses longer. If one group produces fewer or equivalent correct responses but has longer mean Sub-RT, then the interpretation is that the control is more effortful as it took longer to generate the same or a fewer number of items. In contrast, a shorter mean Sub-RT would indicate a faster declining rate of retrieval because a large proportion of the responses were produced early during the trial.

Initiation parameter is the starting point of the logarithmic function that is the value of y when $t = 1$ or $In(t) = 0$ (e.g., initiation parameter for the above mentioned logarithmic function is$y = 4.39 - 1.41\ In(1) = 4.39 - 0 = 4.39$). The initiation parameter indicates the initial linguistic resources or breadth of lexical items available for the initial burst when the trial begins and is largely determined by vocabulary knowledge.

Slope of the plot is determined by the shape of the curve and refers to the rate of the retrieval output as a function of the change in time over 60 seconds. The slope for the above example would be 1.41. It reflects how the linguistic resources are monitored and used over time and is largely determined by executive control. Flatter slope indicates that participants were able to maintain their performance across the response period despite greater lexical interference (e.g., avoiding repetition, searching for words from the already exhausted vocabulary source) towards the end of the trial, reflecting better executive control.

4. Clustering and switching analyses: We closely followed the methods used by Troyer et al.'s (1997). Repetitions were included for the clustering and the switching analyses. Semantic fluency clustering was defined as successively produced words that shared a semantic subcategory. Letter fluency clustering was defined as successively generated words which fulfil any one of the following criteria (Troyer et al., 1997): words that begin with the same first two letters (*stop* and *stone*); words that differ only by a vowel sound regardless of the actual spelling (*son* and *sun*); words that rhyme (*stool* and *school*); or words that are homonyms (*foot:* anatomical part of body, and *foot*: unit of measure). Two variables were generated after clustering the responses: cluster size and number of switches.

Cluster size was calculated beginning with the second word in each cluster. A single word was given a cluster size of zero (e.g., *crocodile*), two words cluster was given a cluster size of one (e.g., *bear, fox* belong to North American animal cluster and cluster size of one), three words cluster was given a cluster size of two (e.g., *rhinoceros*, *hippopotamus*, *deer* belong to African animal cluster and cluster size of two) and so on. Mean cluster size for a trial was calculated by adding the size of each cluster and dividing the total score by the number of clusters.

Number of switches was the number of transitions between clusters. For example, *dog, cat; snake, lizard; horse, cow, goat* contain two switches – before snake and before horse. *Leopard, cheetah; kangaroo, koala bear; robin, sparrow, crow; chimpanzee, orang-utan, baboon* has three switches – before kangaroo, robin and chimpanzee. Similarly, in letter fluency – *fragile, fraught, fray; fan, fat; fly, flower, flute* contain two switches – before *fan* and before *fly*.

Executive control measures

Stroop task (Inhibitory control)

The computerized Stroop Task used in this study was adapted from Scott and Wilshire (2010). It consisted of six colours and their names: red, green, blue, yellow, orange, and purple. The task was divided into two conditions, neutral and incongruent. In the neutral condition, participants named the colour of differently coloured rectangles. A series of 50 coloured rectangles, each in one of the six colours were presented in a random order, such that two successive trials never had the same colour. In the incongruent condition, participants named the font colour of the colour words. A series of 50 colour words were shown one at a time on the screen in a random order, each of which was presented in a colour other than the word's name (e.g., *red* in *green* colour).

The procedure was the same for both conditions. Participants were instructed to name the colour or read the word as quickly and as accurately as possible. Each condition began with six practice trials. Both conditions were completed during a single session with the neutral condition first followed by the incongruent condition. The onset of each stimulus was accompanied by a beep, which allowed latency measurement. All responses were recorded with a digital voice recorder.

Analysis

Accuracy and response times were obtained. The reaction time (RT) analysis was performed after excluding self-corrected and incorrect responses. Using PRAAT, RT for each trial was measured from the onset of the beep to the onset of the naming. Outliers – that is, RTs that were 2.5 standard deviations above or below a participant's mean RT or <250 ms – were removed prior to calculation of the dependent measures. We calculated the Stroop Effect, as the difference between incongruent and neutral conditions (Bialystok et al., 2008; Scott & Wilshire, 2010). Calculation of Stroop Effect can yield similar results even when the interference effects are not similar. For example, for participant 1, RT of 800 ms in the incongruent condition minus a RT of 400 ms in the neutral condition will give a stroop effect of 400 msec. For participant 2, RT of 1200 ms in the incongruent condition minus a reaction time of 800 ms in the neutral condition will also give a Stroop effect of 400 ms. However, the difference score does not take into account overall slowness between the participants. This is a crucial factor in assessing Stroop interference (Green, Grogan, Crinion, Ali, Sutton & Price, 2010). To account for overall speed differences in responses, we calculated Percentage Stroop Ratio (%). The Percentage Stroop ratio (%) was calculated by dividing the Stroop Difference (mean incongruent – mean neutral) by the mean of neutral and incongruent trials, and then multiplied by 100. In the above example, participant 1 and 2 will have a Percentage Stroop ratio (%) of 66.67 and 40, respectively. A smaller Percentage Stroop ratio (%) indicates a

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

better inhibitory control.

$$\text{Percentage Stroop ratio (\%)} = \left[\frac{RT_{INCONGRUENT\ TRIAL} - RT_{NEUTRAL\ TRIAL}}{\frac{RT_{INCONGRUENT\ TRIAL} + RT_{NEUTRAL\ TRIAL}}{2}}\right] * 100$$

Mental-set shifting (Colour-shape switch task)

We adapted Prior and MacWhinney's (2010) colour-shape switch task. Participants had to switch between colour judgement and shape judgement trials. Target stimuli consisted of filled red triangle, red circle, green triangle, and green circle. Participants had to judge the colour or shape of the stimuli based on a cue. There were two types of cues: colour cue (colour gradient) and shape cue (row of small black shapes). If the cue was a colour cue, participants had to judge the colour of the stimulus (red or green) and if the cue was a shape cue, participants had to judge the shape of the stimulus (circle or triangle). The target stimulus appeared at the centre of the screen, followed by the cue that remained on the screen above the target stimulus.

The task was presented via E-Prime (Psychology Software Tools, Pittsburgh, PA). Each trial started with a fixation cross for 500 ms, after which the cue appeared on the screen for 250 ms, 2.8° above the fixation cross, followed by a blank screen for about 300 ms. The targets were red or green circles (2.8° *2.8°) and red or green triangles (2.3°*2.3°). The cue and target remained on the screen until there was a response or for a maximum duration of 2000 ms. This was followed by a blank screen for about 1000 ms before the onset of the next trial. Participants were required to press the key on a computer corresponding to red/green colour or triangle/circle shape.

One half of the trials comprised switch trials, the other half non-switch trials. In the switch trial, a colour stimulus preceded the shape stimulus (colour to shape switch) or a shape preceded the colour stimulus (shape to colour switch). In the non-switch trial, a colour stimulus always preceded another colour stimulus (colour to colour) and a shape stimulus always preceded another shape stimulus (shape to shape). There were 20 practice trials followed by 3 blocks of 48 experimental trials each. There were total 72 switch trials and 72 non-switch trials. Reaction time and accuracy were measured for switch trials and non-switch trials separately. We derived three dependent variables – switch cost for reaction time (SC_{RT}), Percentage switch cost ratio (%), and switch cost for accuracy (SC_{ACC}).

$$SC_{RT} = RT_{SWITCH\ TRIAL} - RT_{NON\text{-}SWITCH\ TRIAL}$$

$$\text{Percentage switch cost ratio (\%)} = \left[\frac{RT_{SWITCH\ TRIAL} - RT_{NON\text{-}SWITCH\ TRIAL}}{\frac{RT_{SWITCH\ TRIAL} + RT_{NON\text{-}SWITCH\ TRIAL}}{2}}\right] * 100$$

$$SC_{Accuracy} = \%Accuracy_{NON\text{-}SWITCH\ TRIAL} - \%Accuracy_{SWITCH\ TRIAL}$$

Smaller switch cost meant participants had a smaller difference (i.e., equivalent performance) between the easier (non-switch trial) and the difficult condition (switch trial). This would suggest efficient shifting ability (Prior & MacWhinney, 2010).

Working memory (backward digit span)

The Wechsler Memory Scale (WMS 3, Wechsler, 1997) was used to measure the backward recall of digit sequences. This is thought to reflect working memory performance (Wilde, Strauss & Tulsky, 2004). Participants were verbally presented an increasingly longer series of digits from 2 to 9, and they were then asked to repeat the sequence of the digits in reverse order. The rate of presentation was one digit per second. The test ended when the participants failed on two consecutive trials at any one span size or when the maximum trial size was reached. The backward digit score was the total number of lists reported correctly in the backward digit span test.

As could be seen in Table 4, the two groups differed significantly on Percentage Stroop ratio (%), Percentage switch cost ratio (%), and switch cost accuracy. Although, bilinguals were overall slower in the Stroop task but there was no difference on the Stroop difference measure. However, when we accounted for overall speed difference, bilinguals demonstrated smaller Percentage Stroop ratio (%) which is indicative of better inhibitory control. Bilinguals also showed a smaller Percentage switch cost ratio (%) and a smaller switch cost accuracy suggestive of superior shifting ability.

Statistical analysis

All verbal fluency measures were normally distributed. To arrive at the mean scores for each measure, the three trials were averaged in each condition; for semantic fluency, animals, fruits and vegetables, and clothing were averaged; for letter fluency *F, A,* and *S* trials were averaged. A two-way ANOVA repeated measure was used on the following measures: number of CR, 1st-RT, Sub-RT, cluster size, and number of switches. In the design, Group (Bilingual, Monolingual) was treated as a between-subject factor, and Condition (Semantic, Letter) was treated as within-subject factor. Tukey's post-hoc tests were applied for significant interaction effects at $p \leq 0.05$. Independent sample t-tests were performed for FDS, initiation parameter and slope for semantic and letter fluency conditions with Group as the between-subject factor. To examine the relationship between the executive control measures and verbal fluency measures, correlations were performed separately for each group.

Results

The mean and standard deviation values for the verbal fluency variables for Group (Bilinguals and Monolingual) and Condition (Semantic and Letter) averaged across participants are presented in Table 5 (standard deviation reflects between-subject variation). The results of the statistical tests are provided in Table 5 as well. Findings from the correlation analyses between the executive control measures and verbal fluency variables for each group are presented in Table 6. Findings for Group differences are presented first, followed by the findings on the relationship of executive control measures and verbal fluency variables. The authors are happy to share anonymized item level time-stamped verbal fluency data with interested readers.

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

Table 3. Mean (M), standard deviations (SD), and statistical results of bilinguals' subjective language profile

	Bengali		English		
Measures	*M*	*SD*	*M*	*SD*	Statistical results
Language acquisition history[7]	14.3[1]	2.6	2.5[1]	2.3	**$t(24) = 14.9$, $p < .001$**
Language of instruction[7]	5.3[2]	1.9	5.9[2]	2.4	$t(24) = -.6$, $p = .53$
Self-rated language proficiency[7]					
Speaking	6.3[3]	0.7	6.4[3]	0.6	$t(24) = -.1$, $p = .91$
Comprehension	6.7[3]	0.4	6.6[3]	0.5	$t(24) = .7$, $p = .50$
Reading	6.7[3]	0.6	6.7[3]	0.5	$t(24) = -.3$, $p = .80$
Writing	6[3]	0.9	6.4[3]	0.6	$t(24) = -1.6$, $p = .13$
Language use[7]	12[4]	1.9	20.6[4]	2	**$t(24) = -14.2$, $p < .001$**
Language dominance[8]	17[5]	3.2	18.2[5]	3.6	$t(24) = -.9$, $p = .37$
Language switching habit[9]	8.7[6]	1.1	7.7[6]	2.1	**$t(24) = -2.3$, $p = .03$**

[1] – maximum possible score was 16, greater score in one language means greater immersion in that language during childhood; [2] – maximum possible score was 9, greater score in one language means greater number of years of education in that language; [3] – on a scale of one to seven (1 = no proficiency, 7 = native like proficiency), greater score in language means greater proficiency in that language; [4] – maximum possible score was 25, greater score in one language means greater use of that language in daily life; [5] – maximum possible score was 31, dominant language is the language which obtains a greater score than the other language; [6] – maximum score possible was 12, greater score in one language means greater switch from that language to the other language; [7] – adapted from Muñoz, Marquardt and Copeland, 1999; [8] – adapted from Dunn and Fox Tree, 2009; [9] – adapted from Rodriguez-Fornells et al., 2012.

Table 4. Means (M), standard deviations (SD), and statistical results of executive control measures

	Bilingual		Monolingual		
Measures	*M*	*SD*	*M*	*SD*	Statistical results
Stroop task[1]					
Percentage Stroop ratio (%)[2]	22.8	8.5	30.7	10.2	**$t(48) = -2.9$, $p = .005$**
Stroop difference (RT)[3]	169.7	64.3	194.9	63	$t(48) = -1.4$, $p = .17$
Stroop incongruent (RT)	835.8	163.7	743	145.1	**$t(48) = 2.1$, $p = .04$**
Stroop neutral (RT)	666.1	146.3	548.1	135.6	**$t(48) = 2.9$, $p = .005$**
Colour-shape switch task[4]					
Percentage switch cost ratio (%)[5]	20	11.5	27.1	11.8	**$t(48) = -2.1$, $p = .04$**
Switch cost (RT)[6]	242.1	154.1	252.5	129.8	$U^9 = 289.5$, $p = .65$
Switch trial (RT)	1330.8	280.9	1035.9	127	**$U^9 = 111$, $p < .001$**
Non-switch trial (RT)	1088.7	242.1	783.4	203.2	**$t(48) = 5.6$, $p < .001$**
Switch cost (accuracy)[7]	0.02	3.3	2.8	4.7	**$t(48) = -2.4$, $p = .02$**
Switch trial (accuracy)	95.2	4.2	91.5	5.7	**$U^9 = 188$, $p = .01$**
Non-switch trial (accuracy)	95.2	4.4	95.2	3.4	$U^9 = 293.5$, $p = .71$
Digit span test[8]					
Backward digit span	6.1	1	5.6	0.9	$U^9 = 226$, $p = .08$

[1] – Stroop task adapted from Scott and Wilshire, 2010; [2] – Percentage Stroop ratio (%) : smaller Percentage Stroop ratio indicates better inhibitory control; [3] – Stroop difference = Incongruent trial mean RT - Neutral trial mean RT; [4] – adapted from Prior and MacWhinney, 2010; [5] – Percentage switch cost ratio (%) : smaller Percentage switch cost ratio indicates better shifting ability; [6] – Switch cost (RT) = Switch trial mean RT - Non-switch trial mean RT; [7] – Switch cost (accuracy) = Non-switch trial mean accuracy - Switch trial mean accuracy; [8] – Digit span test (Wechsler, 1997); [9] – Mann-Whitney U test.

Group differences in verbal fluency performance

Differences between the bilinguals and monolinguals were observed either as a main effect of Group or as an interaction of Group X Condition for CR, FDS, Sub-RT, slope for letter fluency, and cluster size. There were no group differences in 1st-RT, initiation parameters for either semantic or letter fluency, slope for semantic fluency, and number of switches.

The CR showed a main effect of Condition (Semantic: $M = 20.6$, $SD = 3.4$; Letter: $M = 16.8$, $SD = 3.5$) and a significant interaction of Group X Condition (see Figure 1a). Post-hoc analysis of the interaction revealed that there was no significant difference between the groups for semantic condition ($p > .05$). However, bilinguals produced significantly greater number of CR in the letter fluency compared to monolinguals [$t(48) = 1.98$, $p = .05$, d = .53]. For FDS, bilinguals showed significantly smaller FDS

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

Table 5. Means (M), standard deviations (SD), and statistical results of the dependent measures by group (Bilingual, Monolingual) and fluency (Semantic, Letter) conditions

	Bilingual		Monolingual		Total		Statistical results		
Measures	*M*	*SD*	*M*	*SD*	*M*	*SD*	Group	Condition	Group*Condition
CR[1]	18.8	3	18.7	3	18.7	3	$F(1,48) = .01$, $p = .91$, $\eta_p^2 = .001$	**$F(1,48) = 70.2$, $p < .001$, $\eta_p^2 = .59$**	**$F(1,48) = 16.3$, $p < .001$, $\eta_p^2 = .25$**
Semantic	19.8	3.1	21.5	3.5	20.6	3.4			
Letter	17.8	3.5	15.8	3.5	16.8	3.6			
FDS[2]	0.10	0.12	0.26	0.16	0.18	0.16	**$t(48) = -3.9$, $p < .001$, $d = 1.1$**		
1st-RT	1.2	0.5	1.2	0.5	1.2	0.5	$F(1,48) = .01$, $p = .92$, $\eta_p^2 = .001$	$F(1,48) = .18$, $p = .67$, $\eta_p^2 = .004$	$F(1,48) = .34$, $p = .56$, $\eta_p^2 = .007$
Semantic	1.3	0.6	1.2	0.7	1.2	0.6			
Letter	1.2	0.8	1.2	0.7	1.2	0.7			
Sub-RT	23.9	1.5	22.5	1.5	23.2	1.3	**$F(1,48) = 11.4$, $p = .001$, $\eta_p^2 = .19$**	**$F(1,48) = 83.8$, $p < .001$, $\eta_p^2 = .64$**	$F(1,48) = 2.4$, $p = .13$, $\eta_p^2 = .05$
Semantic	21.6	2.4	20.9	2.3	21.2	2.4			
Letter	26.2	1.5	24.1	1.9	25.1	2			
Initiation_semantic	3.9	0.6	4.3	0.9	4.1	0.8	$t(48) = -1.9$, $p = .06$, $d = .54$		
Initiation_letter	2.4	0.5	2.5	0.6	2.5	0.5	$t(48) = -.46$, $p = .64$, $d = .13$		
Slope_semantic	1.2	0.3	1.3	0.5	1.3	0.3	$t(48) = -.81$, $p = .42$, $d = .22$		
Slope_letter	0.5	0.2	0.7	0.2	0.6	0.2	**$t(48) = -2.7$, $p = .008$, $d = .76$**		
Cluster size	0.6	0.2	0.5	0.2	0.6	0.1	$F(1,48) = .4$, $p = .54$, $\eta_p^2 = .008$	**$F(1,48) = 62.1$, $p < .001$, $\eta_p^2 = .56$**	**$F(1,48) = 8.1$, $p = .007$, $\eta_p^2 = .14$**
Semantic	0.7	0.2	0.8	0.2	0.7	0.2			
Letter	0.5	0.2	0.3	0.2	0.4	0.2			
Number of switches	11.7	2.4	12.2	2.1	11.9	2.1	$F(1,48) = .7$, $p = .41$, $\eta_p^2 = .01$	$F(1,48) = 2.9$, $p = .10$, $\eta_p^2 = .06$	$F(1,48) = 1.2$, $p = .27$, $\eta_p^2 = .02$
Semantic	11.9	2.3	12.8	2.1	12.3	2.2			
Letter	11.6	2.8	11.6	2.7	11.6	2.7			

[1] – Number of Correct Responses; [2] – Fluency Difference Score.

Table 6. Correlation coefficients amongst the executive control measures and the verbal fluency measures

Executive control measures		Fluency variables							
		CR	FDS	1st RT	Sub-RT	Initiation	Slope	Cluster size	Number of switches
					Bilinguals ($N = 25$)				
Percentage Stroop ratio (%)									
	rs^1	−0.1	0.30	0.21	0.05	0.33	**0.40***	0.01	−0.13
	p-value	0.62	0.14	0.30	0.81	0.10	**0.01**	0.96	0.52
Percentage Switch cost ratio (%)									
	rs^1	−0.32	0.34	0.11	0.26	0.09	0.36	0.02	−0.31
	p-value	0.11	0.10	0.61	0.21	0.68	0.08	0.91	0.13
Switch cost (RT)									
	rs^2	−0.36	0.29	0.26	0.12	0.02	0.33	−0.08	−0.29
	p-value	0.08	0.15	0.21	0.56	0.93	0.11	0.72	0.16
Backward digit span									
	rs^2	0.23	0.21	0.22	0.27	0.18	0.03	0.02	0.16
	p-value	0.28	0.32	0.28	0.2	0.39	0.87	0.94	0.44
					Monolinguals ($N = 25$)				
Percentage Stroop ratio (%)									
	rs^1	−0.08	−0.14	−0.21	−0.16	0.16	0.20	−0.39	0.16
	p-value	0.69	0.51	0.30	0.44	0.43	0.33	0.06	0.45
Percentage Switch cost ratio (%)									
	rs^1	0.04	−0.25	−0.12	−0.08	0.03	0.12	−0.06	−0.10
	p-value	0.89	0.22	0.55	0.72	0.89	0.58	0.78	0.64
Switch cost (RT)									
	rs^2	0.01	−0.30	−0.05	−0.01	−0.01	0.16	0.13	−0.19
	p-value	0.99	0.14	0.80	0.98	0.98	0.44	0.54	0.38
Backward digit span									
	rs^2	−0.03	−0.03	−0.09	0.21	−0.09	−0.22	−0.01	0.03
	p-value	0.87	0.89	0.67	0.31	0.68	0.28	0.97	0.89

[1] – Pearson's correlation coefficient; [2] – Spearman's rho; *$p \leq .05$.

(Bilingual: $M = .12$, $SD = .15$; Monolingual: $M = .26$, $SD = .16$; see Figure 1b). Sub-RT showed a significant main effect of Group, with bilingual demonstrating longer Sub-RT (Bilingual: $M = 23.9$, $SD = 1.5$; Monolingual: $M = 22.7$, $SD = 1.5$). Cluster size showed a main effect of Condition (Semantic: $M = .7$, $SD = .2$; Letter: $M = .4$, $SD = .2$) and an interaction of Group X Condition (see Figure 1c). Post-hoc analyses revealed that bilinguals produced significantly larger cluster than the monolinguals on the letter fluency condition [$t(48) = 2.3$, $p = .02$, d = .66], however, cluster size was comparable between the bilinguals and monolinguals on the semantic fluency condition [$t(48) = -1.4$, $p = .17$, d = .39].

Figure 2 represents the time-course of the CR by the group for the semantic (Figure 2a) and letter fluency (Figure 2b). The only significant difference between the groups was for the slope for bilinguals in letter fluency. Bilinguals demonstrated a significantly flatter slope than the monolingual in letter fluency. This would suggest that bilinguals were able to maintain their output productivity throughout the 1-minute, especially in the difficult condition (i.e., letter fluency).

Verbal fluency performance and executive control measures

Table 6 presents the correlation coefficients amongst the verbal fluency variables and executive control measures for bilinguals and monolinguals. Bilinguals showed a significant correlation between Percentage Stroop ratio (%) and slope (positive, see Figure 3). Bilinguals with a smaller Stroop ratio illustrated a flatter slope indicating that those with better inhibitory control could maintain their performance throughout the 1-minute of the verbal fluency task. Overall bilinguals showed smaller Percentage Stroop ratio (%) compared to the monolinguals.

Discussion

This research set out to determine group differences in verbal fluency performance between a group of relatively homogeneous Bengali–English bilinguals with English speaking monolinguals, as well as identify the executive control measures that contribute to the performance difference between them. We used a wide

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

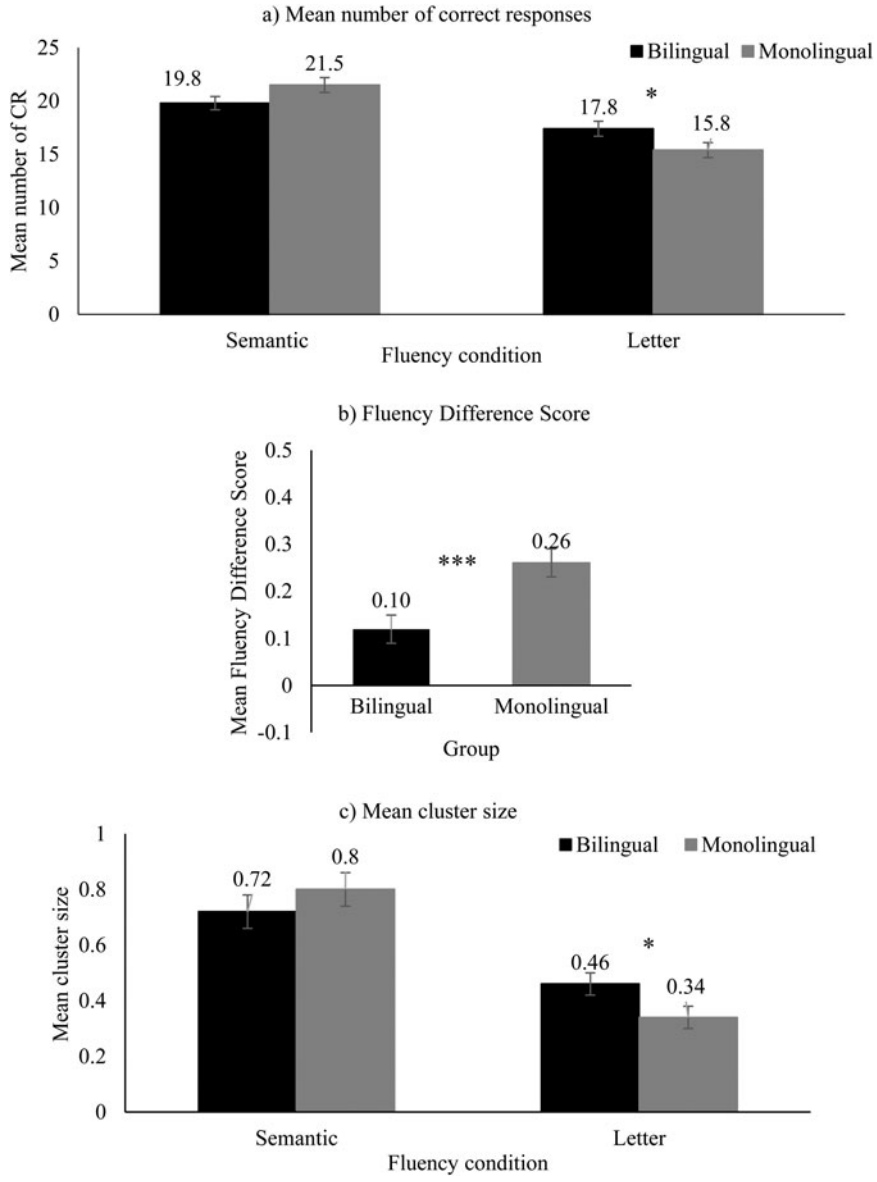


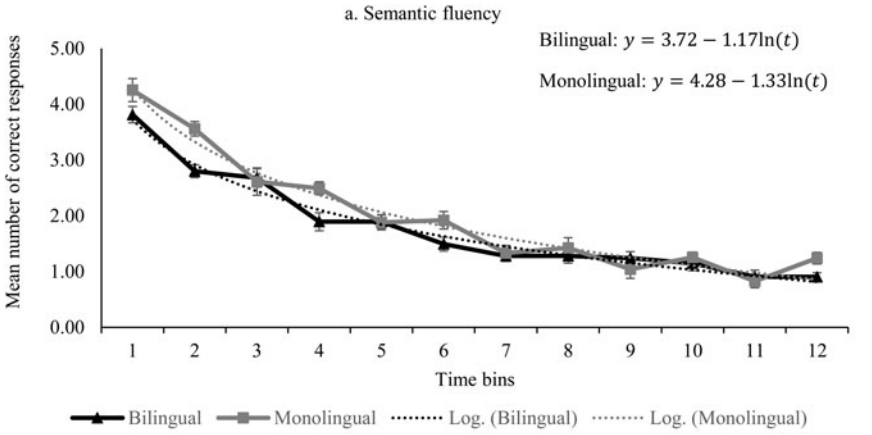
Fig. 1. Verbal fluency variables which revealed significant differences between monolinguals and bilinguals: a) Mean number of correct responses (CR) (top panel); b) Mean Fluency Difference Score (FDS) (middle panel); c) Mean cluster size. Error bars represent standard error of the means; $^{*}p \leq .05$, $^{***}p \leq .001$.

range of measures – CR, FDS, 1st-RT, Sub-RT, initiation, slope, clustering and switching – to characterize the linguistic and executive control components of the participants' verbal fluency performance. These measures are thought to differentially contribute to the linguistic and executive components of verbal fluency task. In addition, we measured executive control in the domains of inhibition, switching, and working memory, and linked the verbal fluency performance to the executive measures.

To summarize the main findings, compared to monolinguals, bilinguals showed differences in both the linguistic (letter fluency: number of CR, cluster size) and executive control (FDS, Sub-RT, slope and number of switches in letter fluency) domains of the verbal fluency task as identified and indicated on Table 7. Although overall there was no significant difference between the two groups on CR, there was an interaction with the type of fluency task. Bilinguals and monolinguals performed similarly on semantic fluency; whilst bilinguals outperformed the monolinguals on letter fluency. The finding that there were no differences regarding CR between the vocabulary matched two groups is consistent with the findings observed in the literature (Bialystok et al., 2008; Luo et al., 2010; Paap et al., 2017; Portocarrero, Burright & Donovick, 2007; Rosselli et al., 2000).

Our findings show that bilinguals perform better than monolinguals in the letter fluency task, which is thought to be more demanding on executive control. This is shown in the following key findings: 1) bilinguals demonstrated significantly smaller FDS than monolinguals, which have claimed to reflect superior executive control; 2) bilinguals demonstrated significantly longer Sub-RT with higher mean number of correct responses in the letter fluency and a flatter slope on letter fluency, which could be attributed to superior executive control. These findings suggest that our bilinguals demonstrate superior executive control abilities which are helping them to perform better (in terms of lower FDS, flatter slope) for a difficult fluency condition (i.e., letter fluency). As discussed in the introduction, longer Sub-RT can be either due to smaller vocabulary or superior executive control abilities of bilinguals compared to monolinguals (Luo et al., 2010). Luo et al. (2010) have postulated that the superior executive control would result in a slower decline in retrieval speed or longer Sub-RT for bilinguals in combination with a higher and or

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press



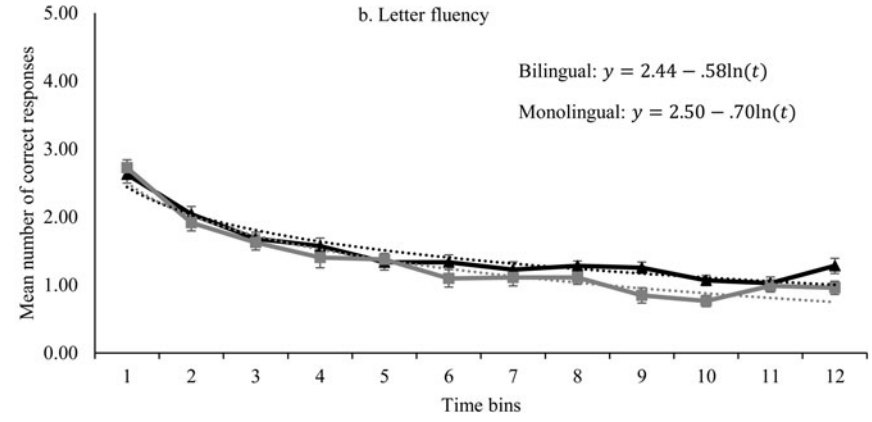


Fig. 2. Between-group comparison of number of correct responses (CR) produced as a function of 5-sec time intervals in: a) semantic (top panel) and b) letter fluency (bottom panel). Best-fit lines are logarithmic functions. Error bars represent standard error of the means.

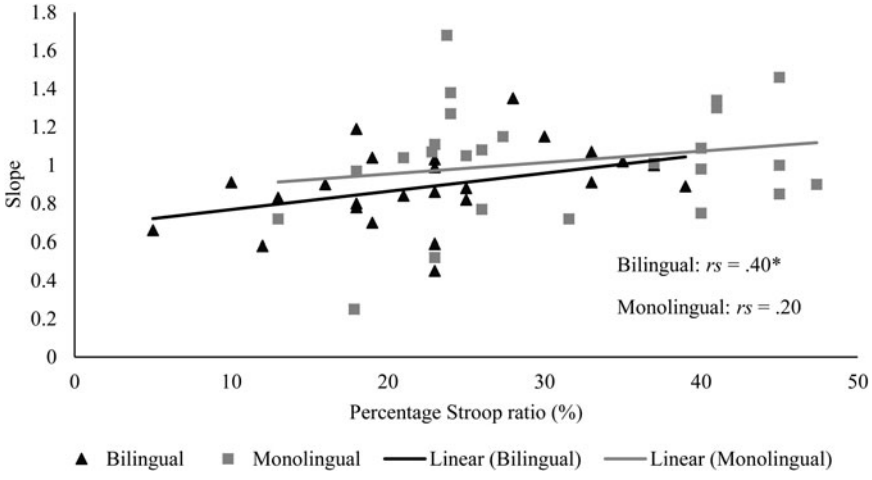


Fig. 3. Correlation plots for the significant correlations between Percentage Stroop ratio (%) and slope of verbal fluency; *rs* represents Pearson's correlation coefficient; $^{*}p \leq .05$.

equal number of CR and flatter slope than monolinguals. Since our groups were matched on vocabulary and we do not find any significant difference between the two groups on the initiation parameter (which is a measure of initial linguistic resources), it would be reasonable to conclude that the bilinguals' performance would be indicative of superior executive control (Friesen et al., 2015; Luo et al., 2010). Overall, equivalent performance on the vocabulary test, longer Sub-RT, and better performance on the letter fluency condition (higher CR, smaller FDS, flatter slope, and larger cluster size) for bilinguals compared to monolinguals suggest a bilingual advantage in the verbal fluency task when there is a higher demand for the controlled executive processing skills.

On the qualitative measures, we expected vocabulary-matched bilinguals to produce equal cluster size, which utilizes more of the linguistic components and a larger number of switches, which requires efficient executive control mechanism. However, we found that bilinguals produced a larger cluster size in the letter fluency condition. This could be due to a strategy to bolster their performance in letter fluency. Greater number of CR with larger cluster size in letter fluency in bilinguals could be a strategy that allowed them to sustain production in a more demanding condition. The lack of a difference in switching is surprising as a switching measure is supposed to tap into the executive control components of the verbal fluency task. We expected bilinguals to

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

Table 7. Results of the current study in the context of verbal fluency measures and to their linguistic and executive control components

	Processes		Bilingual (B) vs. Monolingual (M)	
Parameters	Linguistic	Executive	Differences	Correlation with executive control
Quantitative analysis				
Number of CR[1]	√	√	**Yes**, B > M (letter)	No
FDS[2]		√	**Yes**, B < M	No
Time-course analysis				
1st RT	√		No, B = M	No
Sub-RT		√	**Yes**, B > M	No
Initiation parameter	√		No, B = M	No
Slope		√	**Yes**, B > M (letter)	**Yes**, (+) with Stroop ratio % for B
Qualitative analysis				
Cluster size	√		**Yes**, B > M (letter)	No
Number of switches		√	No, B = M	No

[1] – Correct Responses; [2] – Fluency Difference Score; **Yes** – significant findings, No – not significant findings.

switch more compared to monolinguals. However, no difference between the groups on switching indicates bilinguals may not use switching as a strategy to facilitate their performance in the verbal fluency task.

On the executive control measures, we found bilinguals outperformed monolinguals on the inhibitory control measure (smaller Percentage Stroop ratio), and mental set-shifting measure (smaller Percentage switch cost ratio and smaller switch cost accuracy). However, both groups performed similarly on the working memory measure (backward digit span). An advantage in inhibitory control for bilingual participants is in line with the literature (Bialystok et al., 2004; Bialystok et al., 2008; Emmorey et al., 2008). For the mental-set shifting task, we measured Percentage switch cost ratio (%) to account for the overall speed difference between the two group. We found bilinguals to show advantage in the mental set-shifting measure which is in line with the findings from the previous task switching measures in the literature (Prior & MacWhinney, 2010). However, we did not find any difference between the two groups on the most used dependent variable (switch cost in RT) in the task switching literature. No differences in the switch cost (RT) variable using the colour-shape switch task supports the findings by Paap and his colleagues (Paap & Greenberg, 2013; Paap & Sawi, 2014; Paap et al., 2017). Similarly, having no differences between the two groups on working memory measures is in line with the literature (Bialystok et al., 2008; Luo et al., 2010). Current findings showed that the difference between the two groups on executive control measures might depend on the type of task and the type of dependent variables derived from the task.

Previous studies have suggested the role of executive control measures, especially working memory and inhibitory control in verbal fluency (Luo et al., 2010, Shao et al., 2014). There exists only one study that has directly correlated the executive control measures (updating of working memory and inhibitory control) with verbal fluency measures in healthy monolingual adults (Shao et al., 2014). This is the first study that attempted to establish relationship amongst various executive control measures with measures of verbal fluency comparing bilingual and monolingual healthy adult populations. Results of our correlation analyses showed that verbal fluency slope correlated with inhibitory control (Percentage Stroop ratio) only for the bilingual group (Blumenfeld & Marian, 2011; Prior & Gollan, 2011; see Table 7 and Figure 3). These results support the notion that an executive control advantage helps bilinguals to outperform monolinguals in verbal fluency tasks, especially where executive control demands are higher.

Similar to Shao et al.'s (2014) study, we did not find any correlation between working memory and verbal fluency measures; neither did we find any significant correlation between the mental-set shifting measure and verbal fluency measures. As this was the first study to attempt to establish relationship amongst various executive control and verbal fluency measures, future studies should consider investigating different kinds of tasks within specific domains of executive control to reflect the presumed processes underpinning a verbal fluency task. These lines of research will provide greater insights into the relationship between linguistic and executive control processes during word production.

In conclusion, previous studies comparing healthy monolinguals and bilinguals on verbal fluency tasks have shown mixed results ranging from bilingual advantage (Bialystok et al., 2008; Luo et al., 2010) to disadvantage (Gollan et al., 2002; Paap et al., 2017) to no differences (Paap et al., 2017). However, all these studies have relied ONLY on the number of correct responses as a dependent variable (except Luo et al., 2010). For example, Paap et al. (2017) did not find any difference between bilinguals and monolinguals on the difficult letter fluency condition. The results were inconsistent with the notion that bilinguals' enhanced executive control abilities help them to outperform monolinguals on the more demanding letter fluency condition. Paap et al. also refuted the claim that, compared to semantic fluency, letter fluency requires greater executive control functioning and suggested trying to support this claim by independent and direct tests of executive control abilities. Similarly, Whiteside, Kealey, Semla, Luu, Rice, Basso and Roper (2016) in an exploratory factor analysis study have argued that the contributions of linguistic processes are greater in verbal fluency compared to executive control processes. They found that the number of correct responses in the verbal fluency loaded onto the language factor and not the executive control factor.

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

The present study attempted to address the unresolved issues in the literature by including a wide range of variables and separate measures of executive control abilities. We found that vocabulary-matched healthy bilinguals performed similarly to monolinguals in the semantic fluency task, which is thought to have higher linguistic demands. However, bilinguals outperformed monolinguals in the difficult letter fluency task, which is assumed to have higher executive control demands. Differences between the two groups were observed only on the measures where executive control demands were higher: such as fluency difference score, Sub-RT, slope, and cluster size in the letter fluency. Independent executive control measures (Percentage Stroop ratio) correlated only for the measures (slope) that tapped into the executive control component of the verbal fluency task.

Importantly, only traditional analysis approaches (e.g., number of correct responses) would not have provided a complete picture regarding the relationship between executive control abilities and the performance in verbal fluency task. Both Paap et al. and Whiteside et al.'s study argued against the fact that the letter fluency condition requires greater executive control demands; however, their claim was based on only a number of correct responses as a measure. When a broad range of verbal fluency measures and separate executive control measures were included, we found evidence of executive control involvement in the letter fluency condition. The present study found differences between bilingual and monolingual groups mainly in conditions and measures with higher executive control demands. We found that bilinguals are not at a disadvantage on linguistics measures if they are matched with monolinguals for vocabulary. The present study highlights that, in order to explain advantage, disadvantage, and that there are no differences between bilinguals and monolinguals, it is necessary to use a range of verbal fluency measures and independent executive control measures. Therefore, as recommended by Paap et al. (2017), separate measures of verbal and executive control abilities are necessary to address and explain bilingual advantages and disadvantages in various measures. The extended range of verbal fluency measures used in this study have future implications for understanding the relationship between the executive control impairments and language deficits in a wide range of the clinical population (e.g., aphasia, schizophrenia, dementia, autism spectrum disorders).

Author ORCIDs. Arpita Bose, 0000-0002-0193-5292

Acknowledgements. We are indebted to all our participants for their enthusiasm and time for this research. This research was funded by the Felix Trust, UK.

Supplementary Material. Supplementary material can be found online at https://doi.org/10.1017/S1366728918001098.

References

Abutalebi J and Green DW (2008) Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes* **23**, 557–582.

Bialystok E, Craik FI and Luk G (2008) Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics* **21**, 522–538.

Bialystok E, Craik FI, Klein R and Viswanathan M (2004) Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging* **19**, 290–303.

Boersma P and Weenink D (2015) Praat version 5.4. 08. *Doing phonetics by computer*.

Blumenfeld HK and Marian V (2011) Bilingualism influences inhibitory control in auditory comprehension. *Cognition* **118**, 245–257.

Costa A and Caramazza A (1999) Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language and Cognition* **2**, 231–244.

Crowe SF (1998) Decrease in performance on the verbal fluency test as a function of time: Evaluation in a young healthy sample. *Journal of Clinical and Experimental Neuropsychology* **20**, 391–401.

Delis DC, Kaplan E and Kramer JH (2001) *Delis-Kaplan executive function system (D-KEFS)*. Psychological Corporation.

Dunn AL and Tree JE. F (2009) A quick, gradient bilingual dominance scale. *Bilingualism: Language and Cognition* **12**, 273–289.

Dunn LM (2009) The British picture vocabulary scale. GL Assessment Limited.

Emmorey K, Luk G, Pyers JE and Bialystok E (2008) The source of enhanced cognitive control in bilinguals: Evidence from bimodal bilinguals. *Psychological science* **19**, 1201–1206.

Eng N, Vonk JM, Salzberger M and Yoo N (2018) A cross-linguistic comparison of category and letter fluency: Mandarin and English. *Quarterly Journal of Experimental Psychology*. doi:10.1177/1747021818765997. Published online by Sage Publications, March 28, 2018.

Friesen DC, Luo L, Luk G and Bialystok E (2015) Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience* **30**, 238–250.

Gollan TH, Montoya RI, Fennema-Notestine C and Morris SK (2005) Bilingualism affects picture naming but not picture classification. *Memory & Cognition* **33**(7), 1220–1234.

Gollan TH, Montoya RI and Werner GA (2002) Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology* **16**, 562–576.

Green DW (1998) Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition* **1**, 67–81.

Green DW, Grogan A, Crinion J, Ali N, Sutton C and Price CJ (2010) Language control and parallel recovery of language in individuals with aphasia. *Aphasiology* **24**, 188–209.

Gruenewald PJ and Lockhead GR (1980) The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory* **6**, 225–240.

Hernandez AE and Li P (2007) Age of acquisition: its neural and computational mechanisms. *Psychological Bulletin* **133**, 638–650.

Ivanova I and Costa A (2008) Does bilingualism hamper lexical access in speech production? *Acta psychologica* **127**, 277–288.

Kormi-Nouri R, Moradi AR, Moradi S, Akbari-Zardkhaneh S and Zahedian H (2012) The effect of bilingualism on letter and category fluency tasks in primary school children: Advantage or disadvantage? *Bilingualism: Language and Cognition* **15**, 351–364.

Kousaie S and Phillips NA (2012) Ageing and bilingualism: Absence of a "bilingual advantage" in Stroop interference in a nonimmigrant sample. *The Quarterly Journal of Experimental Psychology* **65**, 356–369.

Luo L, Luk G and Bialystok E (2010) Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition* **114**, 29–41.

Marian V (2008) Bilingual research methods In Altarriba J and Heredia RR (eds), *An Introduction to Bilingualism: Principles and Processes*. Mahawah, NJ: Lawrence Erlbaum, pp. 13–38.

Michael EB and Gollan TH (2005) Being and becoming bilingual. *Handbook of bilingualism: Psycholinguistic Approaches*, 389–407.

Miyake A and Friedman NP (2012) The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current directions in psychological science* **21**, 8–14.

Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A and Wager TD (2000) The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology* **41**, 49–100.

Muñoz ML, Marquardt TP and Copeland G (1999) A comparison of the codeswitching patterns of aphasic and neurologically normal bilingual speakers of English and Spanish. *Brain and Language* **66**, 249–274.

Paap KR and Greenberg ZI (2013) There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology* **66**, 232–258.

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press

Paap KR and Sawi O (2014) Bilingual advantages in executive functioning: problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in Psychology* **5**, 1–15.

Paap KR, Johnson HA and Sawi O (2014) Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *Journal of Cognitive Psychology* **26**, 615–639.

Paap KR, Myuz HA, Anders RT, Bockelman MF, Mikulinsky R and Sawi OM (2017) No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *Journal of Cognitive Psychology* **29**, 89–112.

Portocarrero JS, Burright RG and Donovick PJ (2007) Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology* **22**(3), 415–422.

Prior A and Gollan TH (2011) Good language-switchers are good task-switchers: Evidence from Spanish–English and Mandarin–English bilinguals. *Journal of the International Neuropsychological Society* **17**, 682–691.

Prior A and MacWhinney B (2010) A bilingual advantage in task switching. *Bilingualism: Language and Cognition* **13**, 253–262.

Quick Placement Test. (2001). Oxford: Oxford University Press.

Raven J (2008) *Raven's Standard Progressive Matrices–Plus Version*. London: NCS Pearson.

Rodriguez-Fornells A, Krämer UM, Lorenzo-Seva U, Festman J and Münte TF (2012) Self-assessment of individual differences in language switching. *Frontiers in Psychology* **2**, 1–15.

Rogers CL, Lister JJ, Febo DM, Besing JM and Abrams HB (2006) Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics* **27**(3), 465–485.

Rohrer D, Wixted JT, Salmon DP and Butters N (1995) Retrieval from semantic memory and its implications for Alzheimer's disease. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **21**, 1127–1139.

Rosselli M, Ardila A, Araujo K, Weekes VA, Caracciolo V, Padilla M and Ostrosky-Solí F (2000) Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology* **7**, 17–24.

Sandoval TC, Gollan TH, Ferreira VS and Salmon DP (2010) What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition* **13**, 231–252.

Scott RM and Wilshire CE (2010) Lexical competition for production in a case of nonfluent aphasia: Converging evidence from four different tasks. *Cognitive Neuropsychology* **27**, 505–538.

Shao Z, Janse E, Visser K and Meyer AS (2014) What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology* **5**, 1–10.

Test QP (2001) Oxford: Oxford University Press.

Tröster AI, Fields JA, Testa JA, Paul RH, Blanco CR, Hames KA, Salmon DP and Beatty WW (1998) Cortical and subcortical influences on clustering and switching in the performance of verbal fluency tasks. *Neuropsychologia* **36**(4), 295–304.

Troyer AK, Moscovitch M and Winocur G (1997) Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology* **11**, 138–146.

Troyer AK, Moscovitch M, Winocur G, Alexander MP and Stuss D (1998) Clustering and switching on verbal fluency: the effects of focal frontal-and temporal-lobe lesions. *Neuropsychologia* **36**, 499–504.

Wechsler D (1997) *WMS-III: Wechsler Memory Scale Administration and scoring manual*. Psychological Corporation.

Whiteside DM, Kealey T, Semla M, Luu H, Rice L, Basso MR and Roper B (2016) Verbal fluency: Language or executive function measure? *Applied Neuropsychology: Adult* **23**, 29–34.

Wilde NJ, Strauss E and Tulsky DS (2004) Memory span on the Wechsler scales. *Journal of Clinical and Experimental Neuropsychology* **26**, 539–549.

https://doi.org/10.1017/S1366728918001098 Published online by Cambridge University Press