Measuring Working Memory Capacity with the Letter-Number Sequencing Task: Advantages of Visual Administration

Marta K. Mielicki

Rebecca H. Koppel

Gabriela Valencia

Jennifer Wiley

University of Illinois at Chicago, Chicago,

Illinois, USA

Correspondence

Marta K. Mielicki, University of Illinois at Chicago, 1007 West Harrison Street (M/C 285), Chicago, IL 60607, USA. Email: mmieli2@uic.edu.

Acknowledgements

We are grateful to Sonia Soto for help with data collection and Jennifer Chun for her help with data collection and data coding. This research was supported by grants from the UIC Department of Psychology and the Honors College in support of undergraduate research.

Abstract

Working memory capacity (WMC) plays a major role in many applied contexts, and it is important to be able to accurately measure this construct. The current studies tested whether the modality of administration of the Letter-Number Sequencing task affects performance on the task. The Letter-Number Sequencing task is a WMC measure included as part of the WAIS-III and WAIS-IV test batteries. The task involves hearing a series of letters and digits, and then reporting back the stimuli with the letters in alphabetical order, and digits in ascending numerical order. The task is traditionally administered orally, but recent studies have administered versions of the tasks visually by displaying stimuli on a computer screen. Results suggest that performance differences on the Letter-Number Sequencing task may arise as a function of language background and task administration modality.

Keywords: Working memory capacity, span tasks, bilingualism, modality effects

Measuring Working Memory Capacity with the Letter-Number Sequencing Task: Advantages of Visual Administration

Working memory capacity (WMC) is the ability to concurrently process and store information. WMC is an important construct to study because it plays a major role in many applied contexts. Several researchers have explored the role of WMC in the classroom, specifically looking at WMC deficits in students who experience difficulty in reading and mathematical problem solving (Gathercole et al., 2016; Swanson, 2016). Similarly, individual differences in WMC can affect students' ability to ignore seductive information such as decorative images when learning from multimedia (Sanchez & Wiley, 2006), distraction from music while studying (Christopher & Shelton, 2017), and their ability to monitor their learning accurately (Griffin, Wiley & Thiede, 2008). The role of WMC has been examined in a wide variety of performance domains including driving (Watson et al., 2016), athletics (Furley & Wood, 2016), and simultaneous language interpretation by trainees (Macnamara & Conway, 2016). WMC also plays an important role in legal decision making (Kleider-Offutt, Clevinger, & Bond, 2016) and clinical diagnoses (Hill et al., 2010; Shelton, Elliott, Hill, Calamia, & Gouvier, 2009).

WMC is commonly measured using complex span tasks that feature both storage and processing components (Conway et al., 2005). Complex span tasks measure the ability to control the activation of information in primary memory while engaging in efficient retrieval from secondary memory. The WMC construct can be more generally viewed as a measure of executive or attentional control that represents the ability to maintain information in the face of interference or distraction (Engle, 2002). It is the presence of the processing component that provides the opportunity for interference and distraction and distinguishes complex span tasks from other short-term memory tasks with only storage components (i.e. simple span tasks, Engle et al., 1999). As a measure of attentional control, WMC has emerged as a very important construct in cognitive psychology because it predicts performance on a wide variety of cognitive tasks that require more than just maintenance of information in memory (Conway et al., 2005). These tasks include reading comprehension (Daneman & Carpenter, 1980), problem solving (Wiley & Jarosz, 2012), achievement tests (Engle et al., 1999), and, perhaps most notably, fluid intelligence measures such as the Raven Progressive Matrices (Unsworth & Engle, 2005, 2007). Given the importance of WMC in so many areas of basic and applied cognition, it is an important question whether standard approaches in administering WMC measures might impact the construct validity of WMC scores for a growing portion of the population, namely, bilingual speakers. The current study explores the effects of modality of administration (oral versus visual) on performance on the Letter-Number Sequencing (LNS) task, and tests whether modality effects differ as a function of language background.

The LNS task involves hearing a series of letters and digits in English, and then reporting back the stimuli with the letters in alphabetical order, and digits in ascending numerical order. This task is a standard WMC assessment included as part of the WAIS intelligence test (as a core subtest in WAIS-III and as a supplemental subtest in WAIS-IV; Wechsler, 1997, 2008). It has also been used as a measure of WMC in non-clinical research (Christopher & Shelton, 2017; Macnamara & Conway, 2016). It is traditionally administered orally. As opposed to other tasks in the WAIS such as forward digit span where individuals simply need to store and retrieve the items in the order they are presented, the LNS can be considered a complex span task because of the need to engage in both reordering of the alphanumeric characters (processing), as well as storage and retrieval (Conway et al., 2005; Hill et al., 2010). Shelton et al. (2009) found that the LNS task was highly related to other WMC measures commonly used in laboratory settings.

The current studies explored both traditional oral administration as well as a visually administered version of the LNS task where the stimuli were presented on a computer screen. In the first and second study, a between-subjects design was used to see whether monolinguals and bilinguals perform differently on the LNS task depending on whether the task was administered orally (Experiment 1) or visually (Experiment 2). In the third study, a mixed design was used to test whether administration modality differentially impacted performance for monolinguals and bilinguals.

A related question that has also been of interest to many researchers is whether bilingual individuals may actually experience benefits in executive control. Early research findings supported the idea that prolonged experience managing more than one language could potentially place unique demands on bilingual minds, which could result in cognitive advantages (see Adesope, Lavin, Thompson, & Ungerleider, 2010 for a review). Researchers have explored the idea that bilinguals' need to suppress competition from one language when using another may lead to improved executive control (Bialystok, Craik, Klein, & Viswanathan, 2004), more efficient allocation of executive control resources in the face of conflict, or an advantage in overall response time on tasks that feature competition (Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009; Hilchey & Klein, 2011). Research in other areas of cognition including creative problem solving (Cushen & Wiley, 2011), symbolic thinking (Mielicki, Kacinik, & Wiley, 2017), and learning novel rules (Stocco & Prat, 2014) have also demonstrated bilingual advantages. However, several recent studies and meta-analyses have failed to observe a bilingual advantage on measures of executive control (de Bruin, Treccani, & Della Salla, 2015;

Hilchey, Saint-Aubin, & Klein, 2015; Paap & Greenberg, 2013; Paap, Johnson, & Sawi, 2014; 2015; Paap, Myuz, Anders, Bockelman, Mikulinsky, Sawi, 2017; von Bastian, Souza, & Gade, 2016).

The evidence is also mixed as to whether bilinguals may experience an advantage on measures of WMC specifically, although two recent meta-analyses have found small to medium population effect sizes in favor of a bilingual advantage (Grundy & Timmer, 2017; von Bastian, De Simoni, Kane, Carruth, & Miyake, 2017, November). However, both analyses also found considerable heterogeneity in effect sizes, with many studies included in the meta-analyses demonstrating either no differences in performance between monolinguals and bilinguals or even bilingual disadvantages in performance. Because of this heterogeneity in results, it is an interesting question whether the modality of a WMC task might moderate whether bilingual advantages are found.

Other prior work suggests that bilinguals may be less likely to show advantages, and may even show disadvantages on WMC tasks that rely heavily on verbal or linguistic processing (Engel de Abreu, 2011). Bilinguals tend to have lower vocabulary knowledge in either of their languages than comparable monolinguals (Bialystok, Craik, & Luk, 2008; Bialystok & Luk, 2012). Bilinguals also tend to perform worse on lexical retrieval tasks than monolinguals, even in their dominant language (Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Ivanova & Costa, 2008), and bilinguals may experience more tip-of-the-tongue states than monolinguals (Gollan & Acenas, 2004). Luo, Craik, Moreno, and Bialystok (2013) found that monolinguals outperformed bilinguals on a complex span task that entailed listening to a list of English words and then repeating those words in alphabetical order. However, bilinguals outperformed monolinguals on a spatial complex span task where participants had to remember the order in which blocks were tapped by the experimenter and repeat the taps in reverse order. Bialystok, Craik, and Luk (2008) also observed a bilingual advantage on the same spatial complex span task using blocks, but only for young adult bilinguals and not for older bilinguals.

Another critical difference between the verbal and spatial measures used by Luo et al. (2013), was that they differed in their modality (i.e., oral versus visual task administration). There is evidence to suggest that modality may affect performance on WMC measures. Egeland (2015) administered several WMC tasks from the WAIS-IV (Wechsler, 2008) to a clinical sample, and conducted confirmatory factor analyses to determine whether performance across tasks in different modalities could be best accounted for by the same or different factors. Egeland (2015) found that a model which included separate factors for each modality best fit the data, which suggests that the spatial measures may be tapping into a different construct than the verbal measures. Similarly, Gathercole and colleagues (Gathercole, Durling, Evans, Jeffcock, & Stone, 2008; Jaroslawska, Gathercole, Allen, & Holmes, 2016) have found modality effects on span-like tasks administered to young children, with poorer performance when items were presented orally (which presumably relies on a verbal buffer or the phonological loop) relative to presentation of items via enactment (which may rely on simulated actions stored in a multi-modal episodic buffer).

Other work by Olsthoorn, Andringa, and Hulstijn (2014) has suggested that language background may matter for modality effects. Olsthoorn et al. (2014) administered a complex span task, backwards digit span, to native and non-native speakers of Dutch. Participants received two versions of the task, an oral and a visual version. In the oral version, participants heard the digits in Dutch and in the visual version the digits were presented on a computer screen. Olsthoorn et al. (2014) found that native speakers performed better than non-native speakers overall. Importantly, the groups did not differ in performance on the visual version of the task, but native speakers performed better than non-native speakers on the oral version of the task. The non-native Dutch participants in Olsthoorn et al.’s (2014) study were all late bilinguals who had begun to learn Dutch post-puberty. When Dutch proficiency was entered as a covariate in the analysis, the native-speaker advantage on the oral version of the task was eliminated. Olsthoorn et al. (2014) attributed these findings to native speakers experiencing an *auditory superiority effect*, wherein the oral version of the task is easier than the visual version because it does not require the additional step of retrieving digit names (i.e., linguistic processing) during encoding. Non-native speakers did not experience this auditory superiority effect. Thus, this study suggests that oral administration of WMC tasks may be problematic for some individuals depending on their language background.

Furthermore, work by Sanchez et al. (2010) has suggested that verbal WMC tasks that require processing in a non-native language may not only depress performance on the WMC measure itself, but may also compromise the relation of the measure with performance on general fluid intelligence (gF) tasks such as the Raven’s Advanced Progressive Matrices (RAPM) task. Sanchez et al. (2010) found that non-native English speakers who completed a reading span task in English did worse on the task than native speakers. Moreover, performance on the reading span task did not predict performance on the RAPM task for non-native English speakers, though reading span performance was predictive of RAPM performance for native speakers. Thus, Sanchez et al. (2010) demonstrated that some complex span tasks may not be valid measures of WMC for non-native speakers.

The research summarized above suggests several possibilities for how the modality of a WMC task might differentially affect bilingual and monolingual individuals. If bilingualism confers advantages in WMC (Grundy & Timmer, 2017; von Bastian, De Simoni, Kane, Carruth, & Miyake, 2017, November), then it is possible that a bilingual advantage could be observed for both oral and visual versions of the LNS task. However, there are several possible reasons to predict differences in LNS task performance for monolinguals and bilinguals depending on the modality of task administration. Specifically, performance differences could arise during traditional oral administration of the task. This could be due to bilinguals experiencing additional processing demands when completing the oral version of the task, which could hinder task performance. If both languages are activated during linguistic processing for bilinguals (Kroll & Bialystok, 2013), then oral administration may require suppression of the task-irrelevant language. It is also possible that the oral version of the LNS task may require additional translation processes even for bilinguals who are fluent English speakers. Alternatively, oral administration of the LNS task may be associated with a monolingual advantage because of an auditory superiority effect, as argued by Olsthoorn et al. (2014), whereas bilinguals may not experience this effect. In contrast, because a visual version allows bilinguals to complete the task in whichever language they choose, a bilingual advantage might be observed on the visual version of the LNS task. If performance differences between monolinguals and bilinguals based on task administration modality are observed, this would have implications for the measurement of WMC using LNS task.

Experiment 1

In this experiment, bilingual and monolingual participants completed the standard oral version of LNS. As an improvement over the design of the Olsthoorn et al. (2014) study which selected a native-speaking sample that was higher in non-verbal intelligence than the non-native sample, the groups in this study (and in all studies in this report) were equivalent in scores on a test of college readiness (Math ACT), which addresses the possibility that any differences in the LNS task performance could be attributed to differences in ability between the samples. The participants in all three studies were students at the same U.S. university, and even the bilingual sample largely reported that English was their dominant language. The bilingual samples in all three studies were comprised of self-identified early bilinguals who learned both languages as children. If bilinguals experience an advantage in executive control, then it should translate into better performance on the LNS task. Alternatively, if monolinguals are at an advantage or bilinguals are at a disadvantage due to the oral administration, then bilinguals may perform worse than monolinguals on the oral version of the LNS task.

Method

Participants

A sample of 60 University of Illinois at Chicago undergraduates (34 females, age range 17 to 23) enrolled in introductory psychology courses participated in this experiment in exchange for course credit. Participants were not told that they were being recruited based on their language background for any of these studies. At the end of the study, participants self-identified as either bilingual (n = 30) or monolingual (n = 30). As shown in Table 1, no group differences were found in Math ACT scores, $t < 1$.

Bilingual participants were those who reported prolonged exposure to more than one language before the age of 7, with prolonged exposure defined as both parents speaking a language other than English or attending school taught in a language other than English. Monolingual participants were those who reported being native English speakers with no early prolonged exposure to another language.

For the bilingual sample, 26 participants reported English as their dominant language[1]. The mean age of first exposure to English was 3.28 (*SD* = 3.37) and 1.97 (*SD* = 1.92) for the other language. For the monolingual sample, only five participants reported exposure to a language other than English. The average age of first exposure to another language for these participants was 12.30 (*SD* = 3.77).

Materials

Letter-Number Sequencing Task. The Letter-Number Sequencing subtest from the WAIS-III (Wechsler, 1997) was used in Experiment 1. This task requires participants to listen to a string of alphanumeric characters and repeat the characters back verbally in a specific order (numbers first in ascending order followed by letters in alphabetical order). The stimuli increase in difficulty beginning with 2 alphanumeric characters per string and concluding with 8 characters per string. There are 3 items for each string length for a total of 21 items. There are no repeating letters or numbers within each item (i.e., each character is different within the string). The stimulus items for all experiments are included in the Appendix. Fully correct responses received one point, with a maximum possible score of 21. For comparison across studies, scores are also reported as proportions.

Procedure

After providing informed consent, each participant completed the standard oral version of the LNS task. Stimulus items were presented in the same order for all participants. Stimulus items were read out loud by the experimenter, a native English speaker, at approximately 1 second per character. Participants reported their responses by speaking out loud to the experimenter, and the experimenter marked the string as fully correct or incorrect on a score

[1] The results for performance on the LNS task do not change if the four participants who reported a language other than English as their dominant language are removed from the analysis.

sheet. Responses were not timed. Unlike standard administration, the task was not terminated following three incorrect answers at one string length and all participants attempted all trials. At the end of the study, participants were asked to report their language background and their ACT Math scores.

Results and Discussion

The means and standard deviations for performance on the LNS task are presented in Table 1 by group. Contrary to the prediction of a bilingual advantage in executive control, an independent-samples *t*-test revealed bilingual participants had lower LNS scores relative to monolinguals. No bilingual advantage was observed on the oral version of the LNS task, instead a monolingual advantage, or bilingual disadvantage, was found. The results of Experiment 1 are consistent with those obtained by Luo et al. (2013), who found that bilinguals performed worse than monolinguals on an orally administered verbal measure of WMC. If the bilingual disadvantage found in Experiment 1 is due to oral administration of the LNS task and not due to the verbal nature of the stimuli *per se*, then administering the LNS task visually should reduce or eliminate this effect. Experiment 2 explored this hypothesis.

Experiment 2

In this experiment, monolingual and bilingual participants completed the LNS task on a computer. When the LNS task is administered orally, bilingual participants may need to engage in some translation processes in addition to the cognitive processes required to complete the task. When the LNS task is administered visually and responses are typed rather than spoken, bilingual participants have more flexibility to complete the task in either their L1 or their L2. If the bilingual disadvantage in performance on the LNS task observed in Experiment 1 can be attributed to the mode of task administration, then presenting the task visually on a computer, and collecting responses by participants typing into the computer, should result in similar performance between monolingual and bilingual participants.

Method

Participants

A sample of 60 University of Illinois at Chicago undergraduates (36 females, age range 17 to 21) enrolled in introductory psychology courses participated in this experiment in exchange for course credit. None of these participants had participated in Experiment 1. Participants self-identified as either bilingual (n = 30) or monolingual (n = 30). As shown in Table 1, no group differences were found in Math ACT scores, $t < 1$.

For the bilingual sample, 15 reported English as their first language. The mean age of first exposure to English was 5.23 (*SD* = 2.95) and 1.41 (*SD* = 1.46) for the other language. For the monolingual sample, only three reported learning a language other than English, and the mean age of first exposure to the other language was 13.00 (*SD* = 3.00).

Materials

The same task design for 21 items and scoring were used as in Experiment 1.

Procedure

The procedure was the same as in Experiment 1 except that each participant completed the Letter-Number Sequencing task on a computer. Items were displayed on a computer screen in size 36 Courier New font, and all participants completed all items in the same order. Participants entered their responses by typing them on the keyboard. Each character in each item was presented on screen for 700 ms and there was a 300 ms blank screen between characters. After the termination of the final character in the item, participants were instructed to type the numbers first in ascending order followed by the letters in alphabetical order. Responses were not timed.

Results and Discussion

In contrast to Experiment 1, an independent-samples *t*-test revealed no group differences between monolinguals and bilinguals in scores on the visual version of the LNS task, $t < 1$ (see Table 1). Taken together, the results of Experiments 1 and 2 suggest that the traditional oral administration of the LNS task may advantage monolingual participants or disadvantage bilingual participants. These results are consistent with the perspective that the measurement of WMC may be affected by the modality of administration. However, this evidence was derived by comparing across two studies where different individuals were tested in the oral-administration and visual-administration conditions using different sets of stimuli. There may also have been differences between the two bilingual samples even though they came from the same student population. A mixed design is needed to perform a more rigorous test of whether modality of LNS test administration might impact monolingual and bilingual participants differently.

Experiment 3

In Experiment 3, a mixed design was used to test the hypothesis that modality of LNS task administration might result in differences in performance for monolingual and bilingual participants. Monolingual and bilingual participants were asked to complete the LNS task twice, once with oral and once with visual administration. In addition, while Experiments 1 and 2 used the LNS task based on the WAIS-III, the LNS task has since been further developed in the WAIS-IV. The WAIS-IV version includes the addition of three new 2-character items and 6 new 3-character items in order to improve floor performance (Wechsler, 2008). The LNS task was also revised in an attempt to reduce the number of acoustic errors that occurred due to phonetic similarity of letters and numbers within the strings (i.e., B and C and 3) (Wechsler, 2008, p. 16). Thus, in Experiment 3, a mixed design using stimuli based on the WAIS-IV was used to test whether modality of administration affects monolinguals and bilingual performance on the LNS task differently. In addition to analyzing performance based on the total number of strings that were recalled correctly as in Experiments 1 and 2, in Experiment 3 responses were also scored for the types of errors that participants made on the different versions of the task. The frequency of different types of errors in WMC tasks is thought to relate to individual differences and may reflect different cognitive processing during the task (Jurden, Laipple, & Jones, 1993; Maylor, Vousden, & Brown, 1999; Unsworth & Engle, 2007).

Sattler and Ryan (2009) discuss a classification system for errors that may occur on the LNS (Sattler & Ryan, 2009, p. 117). However, no information regarding the expected frequency for each of these errors is provided and to our knowledge no other research has explored error patterns for the LNS task. This lack of information seems problematic especially since the Interpretation Manual for WAIS-IV states that "unusual responses or response patterns can be very revealing" (Wechsler, 2008, p. 131). Further, Sattler and Ryan (2009) suggest that the types of errors that participants make should also be considered when interpreting the LNS test scores because different types of errors may reflect failure of distinct cognitive processes. Given these concerns, the present study examined the frequency of different error types and whether different patterns of errors would be seen across oral and visual LNS tasks, and across monolingual and bilingual participants.

The findings of the first two studies suggest that bilingual performance on the LNS task is likely to be impacted by modality of administration. Given these findings, one might expect an interaction such that performance will be similar for monolinguals across the oral and visual LNS tasks, yet for bilinguals performance may differ by modality. Further, because in this final experiment modality is manipulated within subjects, and the monolingual and bilingual participants are matched in ability, if such an interaction is obtained, it will provide better insight into whether there might be a bilingual disadvantage or advantage in one of the modalities.

Participants

A sample of 64 University of Illinois at Chicago undergraduates (42 females, age range 18 to 23) enrolled in introductory psychology courses participated in this experiment in exchange for course credit. None of these participants had participated in Experiment 1 or Experiment 2. Participants self-identified as either bilingual ($n = 32$) or monolingual ($n = 32$). Bilinguals ($M = 24.12$, $SD = 4.12$) and monolinguals ($M = 24.07$, $SD = 4.29$) did not differ on Math ACT score, $t < 1$.

For the bilingual sample, 27 participants reported English as their dominant language. The mean age of first exposure to English was 4.61 ($SD = 4.19$) and 5.90 ($SD = 7.72$) for the other language. The mean self-reported proficiency for English on a 10-point scale was 9.27 ($SD = .91$) and for the language other than English was 6.36 ($SD = 3.08$).

For the monolingual sample, only seven participants reported exposure to a language other than English. The mean age of first exposure to English was 1.34 ($SD = 2.06$) and 15.42 ($SD = 6.41$) for the other language. The mean self-reported proficiency in English was 9.86 on a 10-point scale ($SD = .44$), and 1.89 ($SD = 1.95$) for the other language.

Materials

Experiment 3 used stimuli based on the WAIS-IV. To implement the within-subjects design, a second list of items was created (see Appendix) so that each participant could complete one list with oral administration and one list with visual administration[2]. The order of presentation of list (A or B first) and the order of presentation of administration modality (oral or visual first) were counterbalanced across participants. This 2 x 2 design led to 4 versions of the task, and 8 monolinguals and 8 bilinguals completed each version. Within each list, the stimuli increased in difficulty beginning with 2 alphanumeric characters per string and concluding with 8 characters per string, and each list was presented in the same order for all participants. All participants received all trials. Fully correct responses received 1 point. Because the WAIS-IV adds nine items beyond the WAIS-III, the maximum possible score was 30 for each list.

Procedure

The procedure was the same as in Experiment 2, except that each participant completed the LNS task in both modalities. For the oral administration, a recording of the experimenter saying each string at the rate of approximately 1s per character was played on a computer, and participants responded orally. The full response string was also recorded in this study. The visual administration procedure was the same as Experiment 2.

Error Analyses

Because the full response string was recorded in this study, error analyses could be performed for each incorrect trial. Examples of each error type as defined by Sattler and Ryan (2009) are shown in Table 2. Incorrect trials could have more than one error in them, so each trial was coded for the presence or absence of each type of error. When a correct character was omitted, this was coded as an omission error. Auditory errors entailed replacing a character from the original string with a character that was phonetically similar (e.g., replacing a “D” with a “T”). Commission errors entailed adding a character that was not in the original string and was not phonetically similar to a character in the original string. Sequential errors entailed having the

[2] Performance did not differ by list in either the visual or oral conditions, *t*s < 1.16.

correct characters that were reported in an incorrect sequence (i.e. failure to put letters in alphabetical order or digits in numerical order). Characters that were repeated within a response were coded as perseveration errors because there were no repeated characters within any original string. Two independent raters coded all responses for eight participants (480 trials, accounting for 12.5% of the data set) with an interrater reliability for each error type ranging from ICC (1, 2) = .90 to ICC (1, 2) = 1.00. The remainder of the data set was coded by one rater. Errors that did not fit into the predefined categories were quite infrequent (less than 1% of trials). Perseveration errors were also quite rare (less than .5% of trials). Due to their infrequency, these last two error types were not analyzed further.

Results and Discussion

Means for the LNS measures analyzed in Experiment 3 are displayed in Table 3. Total scores (out of 30 items) on the LNS task were analyzed with 2 (monolingual, bilingual) x 2 (oral, visual) mixed ANOVA. The analysis showed no main effect of language background, $F(1,62) = .26$, $p = .61$, but a significant main effect of task modality, $F(1,62) = 4.54$, $p = .04$, $\eta_p^2 = .07$, that was subsumed by a significant interaction between task modality and language background, $F(1, 62) = 5.34$, $p = .02$, $\eta_p^2 = .08$. Follow-up tests of whether modality affected performance within each language group showed that task modality had no effect on total scores for monolinguals, $t(31) = 0.11$, $p = .91$, however bilinguals had higher total scores on the visual version of the LNS task than on the oral version, $t(31) = -2.83$, $p = .008$. Alternatively, the between-subjects comparisons across language groups confirmed that the two groups did not differ in performance on the oral LNS task, $t(62) = 1.21$, $p = .23$, while the bilingual group scored higher than the monolingual group on the visual LNS task, $t(62) = -2.00$, $p = .05$. This pattern of results suggests a possible bilingual advantage in WMC scores when a visual LNS task is used.

The number of trials (out of 30) on which participants committed each type of error are presented in Table 3. Because each incorrect trial could have more than one error type, effects for each error type were tested in a separate 2 (monolingual, bilingual) x 2 (oral, visual) mixed ANOVA. For auditory errors, only a significant main effect for modality was found, $F(1, 62) = 53.93$, $p < .001$, $\eta_p^2 = .47$. Auditory errors were more frequent in the oral LNS task than for the visual LNS task. The main effect for modality was also the only significant effect in the analysis for sequential errors, as sequential errors were more likely on the visual LNS task than the oral LNS task, $F(1, 62) = 4.97$, $p = .03$, $\eta_p^2 = .07$. No significant effects were found in the ANOVA for commission errors. Finally, the analysis of omission errors revealed a significant interaction between task modality and language background, $F(1, 62) = 5.45$, $p = .02$, $\eta_p^2 = .08$. Bilingual participants were more prone to make omission errors in the oral LNS task than monolingual participants, $t(62) = -2.69$, $p = .01$, while no differences were seen across groups in the visual LNS task, $t(62) = .34$, $p = .74$. One interpretation of this pattern of omission errors is that bilingual participants were less able to process the incoming auditory stimuli as efficiently as monolingual participants in the oral LNS task, and this led to less complete initial encoding of the items and thus more omitted items at recall. Regardless of whether this interpretation is correct, the analyses of error types suggests that oral administration may be associated with a greater number of incorrect responses due to phonological errors, whereas visual administration may be associated with ordering errors. Overall these differences in the types of errors that caused incorrect responses suggest differences in the processes that are used during the two versions of the LNS task.

Based on the findings from Sanchez et al. (2010) that WMC tasks requiring high amounts of processing in a non-native language may impact the correlation between WMC measures and measures of gF, data from all three experiments were combined to test whether both versions of the LNS task were predictive of ACT Math across both monolingual and bilingual participants. Prior work has shown that WMC scores and scores on aptitude tests are generally correlated between .30 and .50 (Engle et al., 1999). In the present data, performance on the oral LNS task and ACT Math were correlated at $r = .16$, $p = .10$, and performance on the visual LNS task and ACT Math were correlated at $r = .31$, $p < .001$. These results suggest that standard approaches in administering LNS tasks might impact the construct validity of scores when samples include bilingual speakers.

General Discussion

The results of these experiments suggest that the administration modality of the LNS task alters the processing involved in the task, and impacts bilingual participants and monolingual participants differently. In Experiment 1 bilingual participants scored lower than monolinguals when the LNS task was administered orally, whereas in Experiment 2 no performance differences were found between monolinguals and bilinguals when the LNS task was administered visually. In Experiment 3, which used a mixed design and an updated version of the LNS task, a bilingual advantage was seen specifically with visual administration. An additional analysis combining samples across all three experiments found that the correlation between LNS and Math ACT was compromised when the LNS task was administered orally.

These results are consistent with other work that has demonstrated modality effects for performance on complex span tasks (Egeland, 2015), and modality effects for bilingual participants in particular (Olsthoorn, et al., 2014). However, in contrast to Olsthoorn et al., the pattern of results from Experiment 3 suggests a possible bilingual advantage for visual LNS administration as opposed the monolingual advantage seen on the orally administered WMC task in their study. Although a visual comparison of means from Experiments 1 and 2 suggested that bilinguals were being harmed by the oral administration, using a mixed design in Experiment 3 may have been critical for demonstrating a potential bilingual advantage on the visual LNS task. Additional differences in the LNS tasks may also have contributed to the bilingual advantage that was observed in Experiment 3. The updated LNS task in Experiment 3 reduced the potential for phonological errors and may have altered the extent to which bilinguals were harmed by the oral administration. Further, the additional practice on 2-item and 3-item trials may have helped bilinguals to adapt to the demands of the task. Any of these reasons may help to explain why bilinguals were able to show matched performance to monolinguals on the oral version of the LNS task in Experiment 3.

There are some limitations of the current research. For one, participants across all three experiments were undergraduates, which may restrict the variance that would otherwise be found in the general population. However, even with this population of undergraduates, differences in performance on the LNS task were observed. Another possible limitation is that the participants in this research generally reported English as their dominant language, and the extent to which participants engage in frequent language switching during daily life was not measured. There is some evidence to suggest that the bilingual advantage in executive control may be limited to bilinguals who frequently switch between languages (Prior & Gollan, 2011; Verreyt, Woumans, Vandelanotte, Szmalec, & Duyck, 2016). If the sample of bilinguals used across these experiments did not engage in frequent language switching, this could be a possible reason why a consistent bilingual advantage in LNS performance was not observed. Proponents of the bilingual advantage in executive control could argue that this study did not test for the bilingual advantage under the conditions in which advantages would be most likely.

Another perspective to consider is that the LNS task might measure a "slave system" capacity (such as the auditory or verbal buffer, Egeland, 2015) rather than measuring the central executive or other constructs that represent more overarching executive function (such as WMC, Engle, 2002). Because only a single task was given to each individual, this study cannot provide clarity on this issue. However, it could be argued that differences in performance between monolinguals and bilinguals on the LNS task might not reflect differences in WMC *per se*, but rather differences in a modality-specific "slave system" due to differences in phonological processing.

Although the current studies did not find consistent evidence for a bilingual advantage in performance on the LNS task, the results of this research suggest that administration modality may moderate when bilingual advantages might be observed. These results suggest that administration modality should be carefully considered when administering measures such as the LNS task to bilingual individuals. Thus, the main implications of the current research are for the measurement and, specifically, how the administration of a measure such as the LNS task may impact the measurement of working memory constructs differently for monolinguals and bilinguals.

Since WMC is an important construct implicated in many other complex cognitive abilities (Conway et al., 2005), it is imperative to ensure that commonly used measures of WMC or its components are not biased against certain groups of individuals. Previous research by Luo et al. (2013) and Sanchez et al. (2010) has already demonstrated that complex span tasks which rely heavily on verbal and linguistic processing may pose specific challenges for bilinguals. Beyond the concern with measuring WMC in bilingual samples, the impact of administration modality could also be an important consideration for other vulnerable populations such as elderly participants and young children. Some researchers have already chosen to administer visual versions of tasks similar to the LNS task (Christopher & Shelton, 2017; Emery, Myerson, & Hale, 2007; Macnamara & Conway, 2016). Emery et al. (2007) administered the task to older participants, and, although they do not give a reason for this methodological choice, it is possible that they had concerns about the challenges that oral administration might present for an older adult sample. Similarly, young children may experience modality effects on measures of WMC (Gathercole et al., 2008; Jaroslawska, et al., 2016). Given the importance of WMC in applied contexts, the fact that the LNS task is traditionally administered orally as part of the WAIS, and the importance that is often attached to scores on intelligence tests, it is especially incumbent upon researchers to minimize potential bias on such tasks. The potential bias of WMC measures is particularly problematic in the face of a growing population of individuals who speak more than one language (Fernández & Abe, 2018; Lim et al., 2009; van de Vijver & Tanzer, 2004). The current studies suggest that visual administration may help to avoid the bias that can come from oral administration of WMC tasks.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, *80*, 207-245.

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*, 290-303.

Bialystok, E., Craik, F. I., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 859-873.

Bialystok, E., & Luk, G. (2012). Receptive vocabulary differences in monolingual and bilingual adults. *Bilingualism: Language and Cognition*, *15*, 397-401.

Christopher, E. A., & Shelton J. T. (2017). Individual differences in working memory predict the effect of music on student performance. *Journal of Applied Research in Memory and Cognition, 6*, 167-173.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769-786.

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*, 135-149.

Cushen, P. J., & Wiley, J. (2011). Aha! Voila! Eureka! Bilingualism and insightful problem solving. *Learning and Individual Differences, 21,* 458-462.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450-466.

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science*, *26*, 99-107.

Egeland, J. (2015). Measuring working memory with Digit Span and the Letter-Number Sequencing subtests from the WAIS-IV: Too low manipulation load and risk for underestimating modality effects. *Applied Neuropsychology: Adult*, *22*, 445-451.

Emery, L., Myerson, J., & Hale, S. (2007). Age differences in item manipulation span: The case of letter-number sequencing. *Psychology and Aging*, *22*, 75-83.

Engel de Abreu, P. M. (2011). Working memory in multilingual children: Is there a bilingual effect?. *Memory*, *19*, 529-537.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19-23.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309-331.

Fernández, A. L., & Abe, J. (2018). Bias in cross-cultural neuropsychological testing: Problems and possible solutions. *Culture and Brain, 6*, *1-35.*

Furley, P., & Wood, G. (2016). Working memory, attentional control, and expertise in sports: A review of current literature and directions for future research. *Journal of Applied Research in Memory and Cognition, 5*, 415-425.

Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology, 22*, 1019–1037.

Gathercole, S. E., Woolgar, F., Kievit, R. A., Astle, D., Manly, T., & Holmes, J. (2016). How common are WM deficits in children with difficulties in reading and mathematics? *Journal of Applied Research in Memory and Cognition, 5*, 384-394.

Gollan, T. H., & Acenas, L. A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 246-269.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*, 1220-1234.

Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition*, *36*, 93-103.

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*, 325-340.

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, *18*, 625-658.

Hilchey, M. D., Saint-Aubin, J. & Klein, R. M. (2015). Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks? In J. Schwieter (Ed.) *The Cambridge handbook of bilingual processing* (pp. 586-613). Cambridge: Cambridge University Press.

Hill, B. D., Elliott, E. M., Shelton, J. T., Pella, R. D., O'Jile, J. R., & Gouvier, W. D. (2010). Can we improve the clinical assessment of working memory? An evaluation of the Wechsler

Adult Intelligence Scale–Third Edition using a working memory criterion construct. *Journal of Clinical and Experimental Neuropsychology*, *32*, 315-323.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production?. *Acta Psychologica*, *127*, 277-288.

Jaroslawska, A. J., Gathercole, S. E., Allen, R. J., & Holmes, J. (2016). Following instructions from working memory: Why does action at encoding and recall help? *Memory & Cognition*, *44*, 1183-1191.

Jurden, F. H., Laipple, J. S., & Jones, K. T. (1993). Age differences in memory-span errors: Speed or inhibitory mechanisms?. *The Journal of Genetic Psychology*, *154*, 249-257.

Kleider-Offutt, H., Clevinger, A. M., & Bond, A. D. (2016). Working memory and cognitive load in the legal system: Influences on police shooting decisions, interrogation and jury decisions. *Journal of Applied Research in Memory and Cognition, 5*, 426-433.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*, 497-514.

Lim, Y. Y., Prang, K. H., Cysique, L., Pietrzak, R. H., Snyder, P. J., & Maruff, P. (2009). A method for cross-cultural adaptation of a verbal memory assessment. *Behavior Research Methods*, *41*, 1190-1200.

Luo, L., Craik, F. I., Moreno, S., & Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and Aging*, *28*, 28-34.

Macnamara, B. N., & Conway, A. R. A. (2016). Working memory capacity as a predictor of simultaneous language interpreting performance. *Journal of Applied Research in Memory and Cognition, 5*, 434-444.

Maylor, E. A., Vousden, J. I., & Brown, G. D. (1999). Adult age differences in short-term memory for serial order: Data and a model. *Psychology and Aging*, *14*, 572.

Mielicki, M. K., Kacinik, N. A., & Wiley, J. (2017). Bilingualism and symbolic abstraction: Implications for algebra learning. *Learning and Instruction*, *49*, 242-250.

Olsthoorn, N. M., Andringa, S., & Hulstijn, J. H. (2014). Visual and auditory digit-span performance in native and non-native speakers. *International Journal of Bilingualism*, *18*, 663-673.

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*, 232-258.

Paap, K. R., Johnson, H. A., & Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences?. *Journal of Cognitive Psychology*, *26*, 615-639.

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265-278.

Paap, K. R., Myuz, H. A., Anders, R. T., Bockelman, M. F., Mikulinsky, R., & Sawi, O. M. (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.

Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from Spanish–English and Mandarin–English bilinguals. *Journal of the International Neuropsychological Society*, 17, 1–10

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

Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition*, *34*, 344-355.

Sanchez, C. A., Wiley, J., Miura, T. K., Colflesh, G. J., Ricks, T. R., Jensen, M. S., & Conway, A. R. (2010). Assessing working memory capacity in a non-native language. *Learning and Individual Differences*, *20*, 488-493.

Sattler, J. M., & Ryan, J. J. (2009). *Assessment with the WAIS-IV*. San Diego, CA: Jerome M. Sattler, Publisher, Inc.

Shelton, J. T., Elliott, E. M., Hill, B. D., Calamia, M. R., & Gouvier, W. D. (2009). A comparison of laboratory and clinical working memory tests and their prediction of fluid intelligence. *Intelligence, 37*, 283-293.

Stocco, A., & Prat, C. S. (2014). Bilingualism trains specific brain circuits involved in flexible rule selection and application. *Brain and Language*, *137*, 50-61.

Swanson, H. L. (2016). Word problem solving, working memory and serious math difficulties: Do cognitive strategies really make a difference? *Journal of Applied Research in Memory and Cognition, 5*, 368-383.

Unsworth, N., & Engle, R. W. (2005). Working memory capacity and fluid abilities: Examining the correlation between Operation Span and Raven. *Intelligence*, *33*, 67-81.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*, 1038-1066.

Van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, *54*, 119-135.

Verreyt, N., Woumans, E. V. Y., Vandelanotte, D., Szmalec, A., & Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition*, *19*, 181-190.

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*, 246-258.

von Bastian, C. C., De Simoni, C., Kane, M. J., Carruth, N. P., & Miyake, A. (2017, November). *Does being bilingual entail advantages in working memory? A meta-analysis.* Paper presented at the 58th Annual Meeting of the Psychonomic Society, Vancouver, BC, Canada.

Watson, J. M., Memmott, M. G., Moffitt, C. C., Coleman, J., Turrill, J., Fernández, Á., & Strayer, D. L. (2016). On working memory and a productivity illusion in distracted driving. *Journal of Applied Research in Memory and Cognition, 5*, 445-453.

Wechsler, D. (1997). *WAIS-III administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2008). *WAIS-IV technical and interpretive manual*. San Antonio, TX: NCS Pearson, Inc.

Wiley, J., & Jarosz, A. F. (2012). Working memory capacity, attentional focus, and problem solving. *Current Directions in Psychological Science*, *21*, 258-262.

Table 1.

Means and Standard Deviations for Math ACT and LNS Scores from Experiments 1 and 2

	Monolinguals	Bilinguals	
Task	*M (SD)*	*M (SD)*	*t*-value
ACT Math (E1)	25.31 (3.35)	25.27 (3.53)	0.05
ACT Math (E2)	24.50 (4.24)	24.17 (4.86)	0.28
Oral LNS raw score (out of 21) (E1)	11.10 (2.31)	9.87 (1.70)	2.36*
Visual LNS raw score (out of 21) (E2)	11.30 (2.68)	11.20 (2.25)	0.16
Oral LNS proportion score (E1)	.53 (.11)	.47 (.08)	-
Visual LNS proportion score (E2)	.54 (.13)	.53 (.11)	-

Note: * indicates $p < .05$.

Table 2.

Error Type Examples for Experiment 3

Error type	Correct response	Error response
Omission	2-6-7-D-N-R	2-6-7-D-N
Commission	2-6-7-D-N-R	2-6-7-D-N-S
Auditory	2-6-7-D-N-R	2-6-7-N-R-T
Sequential	2-6-7-D-N-R	2-6-7-D-R-N
Perseveration	2-6-7-D-N-R	2-6-7-D-D-R

Table 3.

Means and Standard Deviations for LNS Measures in Experiment 3

	Monolinguals		Bilinguals	
LNS measure	Oral *M* (*SD*)	Visual *M* (*SD*)	Oral *M* (*SD*)	Visual *M* (*SD*)
LNS raw score (out of 30)	18.34 (2.11)	17.94 (2.72)	17.50 (2.32)	19.03 (2.45)
LNS proportion score	.60 (.07)	.60 (.09)	.58 (.08)	.63 (.11)
Omission error trials	6.31 (2.35)	7.34 (2.54)	7.53 (2.76)	7.19 (2.68)
Commission error trials	4.81 (2.12)	5.53 (2.36)	4.66 (2.52)	4.72 (2.75)
Auditory error trials	5.06 (2.14)	3.03 (1.82)	5.03 (2.89)	2.38 (1.41)
Sequential error trials	1.41 (1.43)	2.09 (2.18)	1.03 (1.03)	1.97 (3.28)