CICSD

Computerized Measures of Verbal Working Memory Performance in Healthy Elderly Participants

Adrienne B. Hancock
The George Washington University, Washington, DC

Leonard L. LaPointe
Julie A. G. Stierwalt
Michelle S. Bourgeois
Tallahassee Memorial HealthCare and Florida State University, Tallahassee, FL

Rolf A. Zwaan
Erasmus University Rotterdam, Rotterdam, The Netherlands

After 30 years, Baddeley and Hitch's (1974) seminal work on a model of working memory continues to serve as a foundation for many empirical studies and commentaries on the mechanism or mechanisms involved in working memory. Working memory is a system for temporarily storing and managing the information required to carry out complex cognitive tasks such as learning, reasoning, and comprehension. Working memory is involved in the selection, initiation, and termination of information processing functions such as encoding, storing, and retrieving data. Figure 1 illustrates the components of Baddeley and Hitch's model: the central executive, which is responsible for planning, integrating, and initiating; the phonological loop, which is responsible for retention and some processing of verbal information; and the visuospatial sketch pad, which is responsible for some processing of visuospatial information.

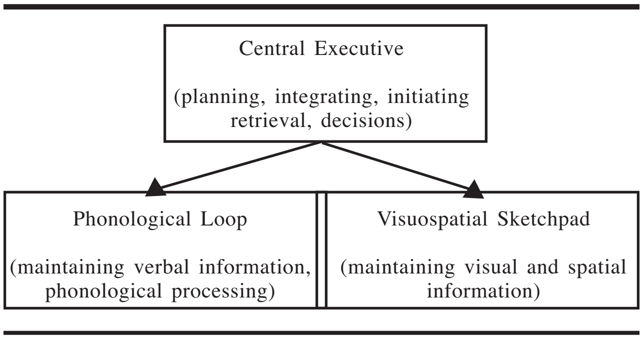
Jonides (1995) summarized evidence implicating working memory involvement in many cognitive tasks, including the language tasks involving comprehension of complex and

ABSTRACT: **Purpose:** Computerized measures of working memory offer advantages of precision, efficiency, and examiner control and are therefore reported more frequently. However, the influence of computer stimulus parameters is still unclear, primarily because they are not uniformly used or reported in published research.
Method: Twenty-four individuals with no reported neurological or serious medical conditions completed a computerized battery of 0-, 1-, and 2-back tasks presented in 2 modalities and 3 levels of interstimulus intervals (ISIs).
Results: Response time (RT) lengthened and accuracy decreased in a 2-back compared to a 1-back task of working memory. Visual or auditory presentation of stimuli generally did not influence RT or accuracy. RT lengthened but accuracy improved as ISI lengthened across 3 levels (800, 1600, and 3200 ms), indicating that participants used the additional time to determine the correct response.
Conclusion: Verbal working memory appears amodal but caution comparing RT across studies using different modalities of stimuli is warranted because RT is affected by perceptual processing time. Influences of ISI are now evident and should be limited by establishing a standard ISI. At minimum, stimulus parameters should be carefully considered and reported in order to advance interpretation of computerized measures.

KEY WORDS: working memory, computerized stimuli, modality, interstimulus interval, response time

1092-5171/07/3402-0073

Figure 1. Baddeley and Hitch's (1974) working memory model.



long sentences, ambiguous words or phrases, and referenced pronouns, as well as judging the relevance or interruption of information (Engle & Conway, 1998). The deleterious effects of a small or taxed working memory in dual-task conditions may appear only during more complex sentence comprehension, indicating that working memory is used on a selective basis. The role of the central executive portion of the working memory model is particularly mysterious. It is pivotal to smooth and efficient cognitive processing but is a challenge to test in isolation.

Although the precise mechanisms and components of working memory continue to be investigated, it is clear that working memory is implicated in language and many forms of higher level cognitive processing. As research accumulates and working memory is further investigated and incorporated into clinical rehabilitation programs, the reliability and other psychometric parameters of measurement tools are crucial to making informed and accurate interpretations of reported working memory assessment strategies. The following discussion explores some of the common methods of measuring working memory and advocates systematic development of computerized measures of working memory and careful exploration of its relationship to language processing.

Traditional Assessment of Working Memory

In working memory tasks, information must be temporarily stored or held available for reference during the manipulation (e.g., calculation, comparison) of information. Several tasks have been developed and used in empirical studies to assess working memory and its various components. Perhaps the most commonly referenced measure of the relationship between working memory and language is the Daneman and Carpenter reading span task (Daneman & Carpenter, 1980). This task requires participants to read sentences of increasing length and remember the last word of each sentence or an unrelated word after the sentence. Researchers who use span measures operate under the premise that the span indicates the amount of processing and storage capacity available while simultaneously using the portions of the same cognitive resources to process the sentence.

Mixed conclusions are reported regarding the use of span measures. The ability to overcome proactive interference from nonfinal words was reported as at least one explanation for differences between ages and variable accuracy of comprehension predictions based on reading span tasks (Lustig, May, & Hasher, 2001; May, Zacks, Hasher, & Malthaup, 1999). Proactive interference is when "older material interferes forward in time with your recollection of the current stimulus" (Ashcroft, 2002, p. 168). In other words, by the end of the task, all of the words you saw since the beginning are still active and are now interfering with the present stimulus. When span stimuli are presented after increasingly longer sentences, the interference is exponential as the task progresses. However, protocols using reading spans beginning with the longest trials result in higher scores than the traditional hierarchical sentence presentation and may be one way to reduce the influence of proactive interference (Lustig et al., 2001; May et al., 1999). Therefore, precisely what working memory span measures (e.g., capacity, processing, both, or other cognitive processes) is inconclusive in the current literature, with nearly as many explanations as there are theoretical perspectives on working memory in general (Towse, Hitch, & Hutton, 2000).

Another task requiring simultaneous storage and processing is the *n*-back task in its many genres. During the presentation of stimuli (e.g., letters, numbers, objects, or locations) in the *n*-back task, the examinee is required to respond to a stimulus when it matches the item seen *n* items ago. For example, in a 1-back task, the participant responds to items that are identical to the item presented one previous to it. For example, participants may see the following numbers presented on the screen one at a time: 5, 8, 3, 6, 6, 2, 9. The participant should respond (press button) when he sees the second 6 because it is the same number as the number he saw 1 back. 3-back tasks require a longer storage or holding duration and more processing because participants must compare each stimulus to the item that was seen or heard three items ago, as well as retain each stimulus for future comparisons and constantly update memory to include only the three most recent items. For example, in the sequence 2, 6, 7, 2, 5, 7, 1, the participant should respond to the second 2 and the second 7 because the same digit was seen 3 back. In this task and other measurement strategies of working memory, temporal duration of this storage or holding time is an important parameter.

Jonides et al. (1997) specified seven processes required for successful *n*-back performance:

- encoding each letter
- storing any letter relevant to a future decision
- rehearsing to keep the items in storage active and ready
- matching when comparing the present item to previous item(s)
- temporal ordering to know which item should be used for comparison to the present item
- inhibiting the oldest item to replace with the newest item in the series
- executing the response.

Some concern is expressed in the literature as to whether the 1-back task is too simple to require the use of working memory. Generally, studies report linear patterns as memory load is increased from 0 to 3 back (e.g., Braver et al., 1997; Schumacher et al., 1996). In contrast, a study using functional MRI (fMRI) reported the dorsolateral prefrontal cortex (DLPFC) to show almost no increase from 0 to 1 back or 2 to 3 back but a large increase from 1 to 2 back, indicating that the DLPFC is not affected by increased memory load until the 2-back task, and that the 2-back task is not much different than the 3-back task (Smith, Jonides, Marshuetz, & Koeppe, 1998). These findings suggest that processes that are used in the 1-back task are more similar to those that are used in the 0-back task, sometimes referred to as choice task, than to the 2- and 3-back tasks that clearly involve working memory. The DLPFC largely contributes the executive control component of working memory (Nystrom et al., 2000). If the DLPFC does not show activity, perhaps the 1-back task indeed does not rely critically on executive control. That would not substantiate the claim that the 1-back task is not working memory, only that it is simple working memory and does not require executive control, which would be indicated by DLPFC activation.

Smith et al. (1998) also reported activation in other working memory areas to increase consistently with memory load. Reviewing the component processes of working memory as listed by Jonides et al. (1997) and discussed above, only the temporal ordering processes, which certainly would require executive control, might be considered unnecessary in the 1-back task. Although the storage and rehearsal demands of the 1-back task are reduced, the participant still performs most of the processes constituting working memory.

Any task that requires simultaneous storage and manipulation of information can be used to assess some component of working memory. Presentation of some traditional tests of working memory, such as the *n*-back task, is easily converted to a computerized version, though there are both advantages and challenges with such a transformation.

Computerized Assessment of Working Memory

Instrumentation of working memory assessment has been refined with technological advances, using precision and reliability available from computerized responses. Although several parameters in computerized assessment are not uniform, the ease and flexibility of programming and administration is appealing for improved precision of clinical and research measurements. Computerized measures of working memory typically require keypad responses to stimuli that are visually presented on a computer screen or aurally presented through a speaker system. Clear directions and practice examples prepare the participant, and criteria for accuracy and speed measures ensure that the participant understands the directions before moving on to the measured responses. Below are a few examples of continuous performance test (CPT) instructions for tests measuring sustained attention, inhibition, and working memory.

- Respond whenever the stimulus X appears.
- Respond when X is followed by another X.
- Respond to X only when X is preceded by A.
- Respond when X is preceded by X-1.
- Respond when anything except X is presented.
- Respond when X is the same as the X presented 3 items previously (*n*-back, specifically 3-back).

Riccio, Reynolds, and Lowe (2001) emphasized that these CPTs are similar but measure very different constructs by varying demands on attention, executive control, and memory. The following discussion will review commonly reported outcome measures of computerized assessments, which typically relate to processing efficiency (response time), accuracy (hit rate, errors of omission and commission), and signal detection theory (sensitivity and response bias), and then turn to the most often manipulated parameters of computerized assessments.

Response time. A common measure of information processing in experiments involving computerized measures is response time (RT), reported in milliseconds (ms). Very complex cognitive tasks require more processing and therefore result in longer RTs when compared to simple tasks. RT reflects total processing from the moment the stimulus is presented until the response key is pressed. Without fMRI data, it is difficult to portion the total RT into distinct types of processing involved (i.e., perception, cognition, motor). Some researchers compare simple reaction time (SRT) to choice reaction time (CRT), deriving the selective attention processing time by simply subtracting SRT, a very simple cognitive task, from the CRT task requiring selective attention. This method dates back to 1868, when Frans Donders proposed "building" complexity to derive time of the additional mental process (as translated in Koster, 1969, and discussed in MacDonald & Meck, 2004; Sternberg, 2001). This requires careful interpretation, as the additive tasks may involve more than one cognitive process or may not use the same time for the original process.

Cautioning the use of Donders' subtraction method, Sternberg (2003) discussed the typically unrealistic assumptions that the original task is accomplished by only one process (which may involve subprocesses) and that the duration attributed to the original task is identical in both measures (e.g., SRT and CRT). More complex cognitive processes, such as working memory, will involve many subprocesses that could complicate Donders' simple derivation formula if comparisons are not appropriate. Another concern, particularly with clinical populations, is the motoric difference between individuals. This somewhat controversial simple subtraction procedure assumes that the motoric contribution to RT will remain constant across tasks for each individual, and that all of the processes (cognitive or motoric response) involved in the SRT task are also involved in the other tasks.

As an alternative, comparisons between the 1-back and 2-back tasks can be made with less fear of violating the "pure insertion" assumption required for the subtraction method. Because the *n*-back tasks primarily differ quantitatively (more load) as opposed to qualitatively (different constructs involved), RT differences between 1-back and 2-back performances reflect the additional load on working memory only. This variation of the subtraction method, which we will term computed working memory (CWM), has been used in neuroimaging analyses of verbal working memory to determine the effects of increasing the load (0-, 1-, 2-, 3-back tasks) in the working memory of healthy individuals and may also be helpful in behavioral studies using RT (Jonides et al., 1997; Smith et al., 1998).

Accuracy measures. In addition to processing speed, researchers are interested in the accuracy of processing. Inevitably, cognitive demands occasionally exceed cognitive ability and errors occur. Measures of accuracy include hit rate, errors of omission (failure to respond to a target), and errors of commission (responding to a foil, also called a false alarm).

If the CPT instructions equally stress the requirements of speed and accuracy, participants are left to develop their own criteria and judgments regarding the relative priorities for speed and accuracy. A conservative participant might have a perfect score yet have a longer RT than someone with a fast RT who commits more errors. Further investigation and other accuracy indices aim to clarify participants' responses for standardization and comparison.

Signal detection theory. Signal detection theory uses a measure of sensitivity (*d´* or d-prime) to describe a participant's response patterns. Although *d´* is the traditional index used in signal detection theory, the nonparametric statistic of sensitivity, *A´*, can be used to avoid possible pitfalls of *d´* and eliminate the necessary precautions required when using collapsed *d´* from averaged data to estimate the average *d´* (Macmillan & Creelman, 1990; Macmillan & Kaplan, 1985). For example, an individual's true hit rate or false alarm rate of one or zero cannot be computed as a *z* score in the *d´* formula. An obstacle of *d´* that is more specific to group data is that if participants' sensitivities or response criteria are disparate, averaged *d´* is not an accurate reflection of performance (Macmillan & Kaplan, 1985). Sensitivity (*d´* or nonparametric *A´*) measures the ability to distinguish between stimulus classes. A good sensitivity score would indicate that the participant discriminated among stimuli and differentiated targets from nontargets, specifically (Riccio et al., 2001). Impulsive responders might have high hit rates and appear quite accurate without considering the rate of commission or false alarms.

A goal of responsible and accurate measurement should be to reduce or eliminate the effects of extraneous variables contributing to the speed and accuracy measures but not stemming from the assessed cognitive process. For example, providing clear instructions for responding quickly and accurately, or setting criteria for practice sessions to be sure the task is comprehended, reduces anxiety influences. In addition to the many measures available from computerized assessment, several factors should be considered when selecting or developing stimuli for the computerized measure of attention or memory. Parameters of stimulus modality and interstimulus interval (ISI) can be manipulated precisely to suit the examiner; however, the effects of manipulating these variables are still not clearly understood.

Modality of stimuli. Stimuli in CPTs can be numbers, letters, words, or particular categories (e.g., animals, real words vs. nonwords). Visual tasks may use geometric forms, colors, pictures, or position. Tests of spatial ability often display stimuli and ask the participant to respond to a certain placement or configuration of items. Most auditory tasks use various tones; some use letters or words.

Of 229 studies using CPTs, indexed by Riccio et al. (2001), only 27 presented at least some auditory stimuli, and even fewer studies directly compared modalities. According to Riccio et al., performance is generally slower and less accurate in tests of auditory (compared to visual) vigilance and may be a more sensitive measure for identifying problems with vigilance and executive control.

However, the influence of stimulus modality specifically on verbal working memory has not been adequately explored. Baddeley and Hitch's (1974) model of working memory suggests that all verbal input would be processed (stored and rehearsed) in the phonological loop, making input modality irrelevant. Based on studies using predominately visual stimuli, neuroimaging data support frontal-parietal involvement in verbal working memory. However, some neuroimaging data and case studies of individuals with brain damage suggest that the inferior parietal cortex and some subcortical structures (e.g., thalamus) may be involved exclusively in auditory–verbal working memory and lend support for a verbal working memory system or at least components of a system that is specific to auditory input (Schumacher et al., 1996; Smith et al., 1998).

Schumacher et al. (1996) used positron emission tomography (PET) and a subtraction method to deduce common areas of brain activation during visual and auditory 3-back tasks. Accuracy was not affected by modality, but RT was significantly increased in the auditory 3-back task, perhaps due to encoding differences. Interestingly, despite the differences on behavioral measures, generally the same neural areas were activated in both modality conditions and few quantitative differences were found; the only differences were in the right supplementary motor area and the inferior portion of Broca's area (more activation in auditory condition). Further calculations between modality and control conditions indicated that Brodman's area (BA) 18, in the occipital lobe, may translate visual representations to phonological, whereas BA 21 and 22 of the temporal lobe are activated during tasks when auditory representations are translated to phonological. Schumacher et al. concluded that verbal working memory is amodal, simply using phonological representations that, one way or another, have been translated from visual or auditory stimuli.

ISI. In addition to the actual stimulus duration, there exists a time interval between each presentation of a stimulus, referred to as the ISI. ISIs may be preset in trial blocks (test generated) or programmed to adjust based on the accuracy of the participant's prior response (participant driven). Empirical studies rarely provide a strong rationale for selecting a particular ISI. It can be difficult to compare results of studies using different ISIs as there is great variation across studies because no standard has been suggested for a short, medium, or long ISI.

Okazaki et al. (2004) reported RT differences between conditions of ISIs of relatively short (800 ms), medium (1500 ms), and long (3000 ms) durations in a cued CPT (i.e., press button when *9* appeared after a *1*). Healthy child and adult participants performed better on the cued CPTs when ISI was longest; RT decreased and accuracy (i.e., hit rate) improved (Okazaki et al., 2004). However, Silverstein, Weinstein, and Turnbull (2004) tested 107 healthy adults and found no difference in RT between 650 ms and 900 ms ISI, but accuracy did improve in the longer ISI condition. Chee, Logan, Schachar, Lindsay, and Wachsmuth (1989) tested 51 hyperactive boys using three levels of ISI (800, 1600, 3200 ms) in a sustained attention task and found the opposite results: Longer ISI increased RT. The disparate results of these studies partly stem from the inconsistent ISI. Without an understanding of ISI's influence and some convergence on a standard methodology, the usefulness of empirical findings is limited.

ISI is particularly important to consider in tasks of working memory, in which participants are required to retain and manipulate information for the duration of the ISI. If the ISI is too brief, participants will not have time to process and respond. If the ISI is longer, the storage demands of working memory will be more taxed. Barch et al. (1997) compared ISIs of 1000 ms and 8000 ms in a CPT of simple working memory (i.e., "Respond to X when it follows A"). The latter ISI is much longer than in most CPTs. Although no behavioral differences in RT or *d´* reached significance in this group of healthy participants, fMRI revealed greater neural activation in the DLPFC, Broca's area, and left posterior parietal area during the long ISI (Barch et al., 1997). Individuals with degeneration or dysfunction in these areas (i.e., people with Parkinson's disease, aphasia, traumatic brain injury) may demonstrate behavioral impairment (as a result of neurological limitations) on CPT measures of working memory. Riccio et al. (2001) summarized the small ISI literature to conclude that additional studies should explore the influence of ISI across various CPTs and populations.

RATIONALE AND STATEMENT OF THE PROBLEM

Computerized measurement of working memory for clinical and theoretical purposes has numerous advantages, but many methodological questions remain. The influence of computerized task and stimulus properties has been explored and documented in some studies of healthy and impaired populations, although few specifically investigate verbal working memory. Several studies have contributed to differentiating types of working memory (e.g., verbal, spatial), but fewer have investigated differences within each type, such as the influence of modality. Although most CPTs present stimuli in the visual modality, investigation of the auditory modality and comparisons between performance in the visual and auditory modalities may prove useful in the investigation of working memory and cognitive function and may guide researchers in further investigation of the executive control processes that comprise the overlap between visual and auditory working memory tasks.

With evidence suggesting that modality and ISI can influence performance, an important research path appears to be to explore these variables and determine whether or not manipulation of these parameters affects derived performance results. Clear understanding of the effect of these assessment variables may contribute to our understanding of working memory and perhaps influence working memory assessment and interventions. In summary, computerized measures of verbal working memory are promising tools, but the influence of important parameters, including stimulus modality and ISI, have not been explored adequately.

Therefore, the purposes of the present study were to compare auditory and visual *n*-back verbal working memory task performance and to explore the effect of ISI on RTs and the accuracy of responses.

The objectives and overall general purposes of this study have generated the following specific hypotheses:

- Performance on verbal working memory tasks will not be significantly affected by modality, especially after implementing a parametric subtraction method to isolate the working memory component.
- Memory load will increase with longer ISIs and in a 2-back compared to a 1-back task. As memory load is increased, RT will lengthen and accuracy will decrease in both auditory and visual modalities of the *n*-back tasks.

METHOD

Participants

Twenty-four adults (17 male, 7 female) with a mean age of 72 years (*SD* = 10.7) and mean educational experience of 16 years (*SD* = 3) participated in the current study on a voluntary basis. Participants were screened for dementia using the Dementia Rating Scale 2 (DRS–2; Jurica, Leitten, & Mattis, 2001). In addition to the presence of dementia, any reported history of psychological disorder, neurological disorder or trauma, or drug or alcohol abuse excluded the individual from this study. Individuals were native speakers of English and were able to name five of five letters presented on a computer screen in the same font, size, resolution, and position as the stimuli in the cognitive battery. Participants were required to have hearing levels characterized by a pure-tone average of 40 dB or less (corrected or uncorrected) in at least one ear, as tested by the examiner using a calibrated portable audiometer. Individuals were allowed to use corrective lenses when necessary. Participants were recruited from the Tallahassee and Clearwater, FL communities (e.g., senior centers, churches, and volunteer groups). All 24 individuals qualified and completed all testing.

Materials

The DRS–2 is a standardized measure of cognitive status for adults (aged 55–105) with cortical impairment, particularly of the degenerative type (Jurica, Leitten, & Mattis, 2001). The test–retest reliability of the DRS–2 is .97, and the DRS–2 is significantly correlated with the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975; $r = .82$) and the Wechsler Adult Intelligence Scale (Wechsler, 1997; $r = .75$) (Jurica et al., 2001). Although the large normative sample consists mostly of Caucasian adults, the majority of participants in the current study were also Caucasian.

The computerized measure of attention and working memory was administered using a laptop computer with a 15-inch monitor and headphones. E-prime (Schneider, Eschman, & Zuccolotto, 2002) was used to customize a computerized assessment battery including the following tasks:

- Visual Simple Response Time 1 (vSRT1): "Respond to any letter you see."
- Auditory Simple Response Time 1 (aSRT1): "Respond to any letter you hear."
- 1-back visual (three tests with varied ISIs: 800, 1600, and 3200 ms)
- 1-back auditory (three tests with varied ISIs: 800, 1600, and 3200 ms)
- 2-back visual (three tests with varied ISIs: 800, 1600, and 3200 ms)
- 2-back auditory (three tests with varied ISIs: 800, 1600, and 3200 ms)
- Visual Simple Response Time 2 (vSRT2): "Respond to any letter you see."
- Auditory Simple Response Time 2 (aSRT2): "Respond to any letter you hear."

Every task included 45 letter stimuli that were selected from a set of 14 letters (all consonants except B, L, M, S, V, W, and Y) (Nystrom et al., 2000; Schumacher et al., 1996). During the visual tasks, black, bold 48-point Times New Roman font letters were displayed in the center of a white screen. To prevent visual matching strategies, letters varied between capital and lower case (Schumacher et al., 1996). Half of the targets in the visual *n*-back tasks were lower case–upper case pairs; the remaining half were matched case pairs. A fixation cross was displayed during each auditory stimulus. The letters presented aurally were varied randomly between male and female voices.

The auditory stimuli were recorded using Multi-Speech model 3700 (Kay Elemetrics, 2002) and were presented at the participant's most comfortable listening level (Lehman & Tompkins, 1998). Duration of the auditory stimuli was between 400 ms to 700 ms; the visual stimuli were set at a constant 500 ms. Based on other studies of *n*-back tasks, target frequency was set at 33%, or 15 out of 45 trials (Schumacher et al., 1996; Wei et al., 2004). ISI for the SRT tasks varied between 800 ms and 1200 ms. Slight variation around the standard 1000 ms was intended to prevent participants from using an anticipation strategy.

Procedure

Individuals were recruited by phone or at community meetings. Testing occurred at the Regional Rehabilitation Center, the First Baptist Church, or the participant's residence. The screening tasks required 30–40 min and the computerized assessment required approximately 45 min, with breaks allowed at the participant's request. All testing environments were quiet and participants were instructed to ignore any distractions (e.g., phone ringing). After hearing an explanation of the experiment and an overview of the testing session, participants were able to ask questions before signing an approved Institutional Review Board consent form.

The examiner explained that the participant would be instructed to press the spacebar using his or her dominant hand when certain stimuli appeared on the screen or were heard through the headphones, and that although the type of target may change, the response was always to press the space bar. The examiner stressed the importance of both speed and accuracy and instructed the participant to try his or her best. A practice session was presented before each type of task. Participants were allowed up to three practice sessions until a passing criterion of 80% accuracy was achieved to ensure that the participant understood the directions and to reduce any task switching effects. All participants were able to achieve criterion within three practice sessions.

For all computerized tests, RT, true positives (hits), omission errors, and false positives (false alarms) were recorded. Conditions were counterbalanced across modality and ISI for the *n*-back tasks. A mandatory comfort break in the middle of the computerized battery was implemented to further reduce any influence of fatigue or boredom.

Variables

For each response, the computer recorded RT in milliseconds as a measure of processing speed. In the current study, participants were informed that both RT and accuracy would be measured; therefore, the response criteria were determined independently and were not necessarily uniform across participants. By using *A´* instead of *d´*, concerns about independent criteria selection were avoided. True and false positives (i.e., hits and false alarms) in each test were recorded and were used to calculate *A´* according to the formulas presented by Donaldson (1992).

Several independent variables were manipulated for investigation:

- cognitive task (SRT, 1-back, 2-back)
- modality (auditory or visual) within all tasks
- ISI (800 ms, 1600 ms, 3200 ms) within 1-back and 2-back tasks.

The *n*-back tasks were selected because they are commonly used tests of working memory, requiring both storage and comparison of stimuli (Jonides et al., 1997). Unlike some working memory tasks, the *n*-back tasks can be included in a computerized measure and require an ISI. The blocked ISI durations of 800, 1600, and 3200 ms were selected based on previous studies investigating the effects of ISI in other populations and include the range of ISIs that are commonly reported (Chee et al., 1989; Okazaki et al., 2004).

RESULTS

Removing Outliers and Deviant Stimuli

Although the process of selecting letters for stimulus items was based on previous research and aimed to reduce ambiguity or confusion (e.g., *m* was excluded because it sounds like *n*), RT and accuracy data were analyzed to determine if a particular letter, presented visually or aurally, influenced performance. Several parameters of each participant response were investigated to uncover stimuli associated with performance deviating from that associated with any other stimuli.

RTs were pooled across participants and ISI (800, 1600, 3200) and were then grouped by letter (14 were included in the testing) and type of presentation (lower case, upper case, male voice, female voice) for each type of test (SRT, 1-back, 2-back). Means and standard deviations appeared to vary across tasks, but until later analysis compares the tests and groups, the coefficient of variation ratio (a ratio of *SD* to *M*) is a useful statistic for comparing across these tasks. Coefficients of variation ranged from .19 to .61, with no obvious deviations. Therefore, no letters were deemed odd or influential based on RT performance.

The above RT analysis investigated the possibility that some stimuli may have required more time for perception, which would alter the total RT. Another possibility is that the participant took too long to determine the stimulus or determined the stimulus to be something else. For example, if the participant was to respond to the presented *p* stimulus, but incorrectly perceived it as *t*, his nonresponse—although correct from his perspective—would be scored as an omission and no RT would be recorded. Because these instances would not be included in the above analysis, an additional analysis of the stimuli using omission rate was performed. The rate of omission (proportion of targets missed) was calculated for each letter in all four types of stimulus presentation for both tests of working memory. A relatively high omission rate is indicative of a stimulus that was perceived incorrectly or was too difficult to perceive in the allotted response time. The *z* stimulus presented aurally with the female voice resulted in the highest omission rates, which approach chance performance, in both 1-back and 2-back tasks. Therefore, all responses to the female *z* stimuli in all tests were removed before analyzing the variables of interest in this study.

Analysis of Dependent Measures

Inclusion of SRT1 and SRT2 was intended to monitor for fatigue or practice effects over the duration of the testing. A paired samples *t* test comparing SRT1 and SRT2 in each modality revealed significant differences between the tasks for both visual ($t = 3.153$, $df = 47$, $p = .003$, two-tailed) and auditory ($t = 3.673$, $df = 47$, $p = .001$, two-tailed) modalities. Despite these unexpected differences, Cohen's effect size was small to medium for comparisons between first and second presentations of SRT for both visual ($d = .46$) and auditory ($d = .53$) modalities. Furthermore, scatter plots and paired samples correlations revealed that the SRT tasks were correlated in both visual ($r = .539$, $p < .000$, $r^2 = .29$, large effect) and auditory ($r = .650$, $p < .000$, $r^2 = .42$, large effect) modalities, providing additional assurance that performances in the identical tasks were related. In visual and auditory modalities of presentation, the mean RT was always shorter during the second presentation of SRT, implying that participants' fatigue did not interfere with performance but rather additional confidence or strategy may have developed through the testing session. SRT1 and SRT2 scores were kept separate in further analysis despite confidence that the statistically significant difference between these two identical tasks was of minimal importance.

Descriptive statistics are reported for the dependent measures of RT and accuracy in all conditions (see Table 1 and Figures 2 and 3). Visual inspection of the data reveals that the RT appears to lengthen as the ISI increases, and the RT appears shorter while *A´* remains high for tests with visual stimuli compared to auditory stimuli. Several statistical analyses were performed to investigate the differences objectively.

Comparison of Four Cognitive Tasks and Modality

Each task presented auditory and visual stimuli, but only the working memory tasks varied ISI level; therefore, levels of ISI in *n*-back tasks were pooled to analyze the influence of modality in all cognitive tasks in a 4 (all tasks) × 2 (modality with pooled ISI) repeated measures analysis of variance (ANOVA; $\alpha = .05$) using the mean RT. Significant main effects were found for comparisons of task, $F(2.1, 96.6) = 329.228$, $p < .000$, $\eta_p^2 = .877$, and modality, $F(1, 46) = 676.512$, $p < .000$, $\eta_p^2 = .936$. Pairwise comparisons with Bonferroni adjustments revealed significant differences between modalities and between all four types of tasks. RT was significantly longer when stimuli were presented aurally and during the 2-back task. Another ANOVA using hit rate instead of RT found statistically significant differences between all four tasks, $F(1.5, 71.2) = 116.707$, $p = .000$, $\eta_p^2 = .717$), but not modalities $F(1, 46) = .827$, $p = .368$, $\eta_p^2 = .018$). Auditory presentation of stimuli increased RT but did not significantly affect hit rate.

Comparison of Two Working Memory Tasks and Modality and ISI Level

Additionally, the influence of both modality and ISI specifically in working memory tasks was evaluated in a repeated measures ANOVA: 2 (*n*-back tasks) × 2 (modality) × 3 (ISI) investigating mean RT. Significant main effects were found for *n*-back tasks, $F(1, 46) = 110.501$, $p < .000$, $\eta_p^2 = .706$); modality, $F(1, 46) = 374.214$, $p < .000$, $\eta_p^2 = .891$); and ISI, $F(1.5, 66.9) = 77.306$, $p < .000$, $\eta_p^2 = .627$). However, two-way interactions between modality and ISI as well as between task and ISI reached significance, and a three-way interaction between modality, task, and ISI was also significant, implicating ISI as influencing the reported effect of task and modality. Pairwise comparisons with

Table 1. Mean response time (RT), hit rate, and sensitivity (A') in four tasks, two modalities, and three levels of interstimulus interval (ISI).

	RT	*Hit rate*	A'
SRT 1 auditory	695	0.99	
SRT 1 visual	345	0.99	
1-back auditory 800 ISI	817	0.91	0.89
1-back auditory 1600 ISI	878	0.92	0.94
1-back auditory 3200 ISI	885	0.96	0.94
1-back visual 800 ISI	553	0.92	0.91
1-back visual 1600 ISI	576	0.94	0.93
1-back visual 3200 ISI	604	0.94	0.92
2-back auditory 800 ISI	860	0.63	0.64
2-back auditory 1600 ISI	1026	0.81	0.78
2-back auditory 3200 ISI	1186	0.80	0.76
2-back visual 800 ISI	690	0.60	0.64
2-back visual 1600 ISI	741	0.72	0.70
2-back visual 3200 ISI	761	0.83	0.78
SRT2 auditory	616	0.99	
SRT 2 visual	320	0.96	

Note. SRT = simply response time. A' cannot be calculated for SRT tests because there are no opportunities for false alarm responses in the SRT test.

Bonferroni adjustments revealed significant differences between the two modalities, between the two *n*-back tasks, and between all three levels of ISI (see Table 2). As in the analysis of all four tasks, RT was significantly longer when stimuli were presented aurally, and RT in the 2-back task was significantly longer than RT in the 1-back task. RT increased significantly with each increase of ISI duration.

This 2 (*n*-back tasks) × 2 (modality) × 3 (ISI) repeated measures ANOVA was repeated using A' to investigate the influences on accuracy of working memory performance. Although the above ANOVA suggested that modality had the largest effect on RT, the ANOVA using A' reported the largest effect to result from the *n*-back task. Only the main effects of *n*-back task, $F(1, 46) = 275.792$, $p < .000$, $\eta_p^2 = .857$, and ISI, $F(2, 92) = 36.160$, $p < .000$, $\eta_p^2 = .440$, are significant when using A'. Interactions between *n*-back task and modality, $F(1, 46) = 7.44$, $p = .009$, $\eta_p^2 = .139$, and between *n*-back task and ISI, $F(2, 86) = 12.383$, $p < .000$, $\eta_p^2 = .212$, also were significant, this time implicating level of *n*-back as the influential variable. Table 3 displays the results of pairwise comparisons performed as a result of this ANOVA.

Isolating Working Memory Processing

In addition to the ANOVA described above using mean RT in 1-back and 2-back tasks separately, a 2 (modality) × 3 (ISI) repeated measures ANOVA ($\alpha = .05$) used CWM (discussed above as an acceptable form of a subtraction method: subtracting 1-back RT from 2-back RT) to explore the effects of modality or duration of ISI with less concern about possible influences of different cognitive processes, and controlling for individual differences in motoric

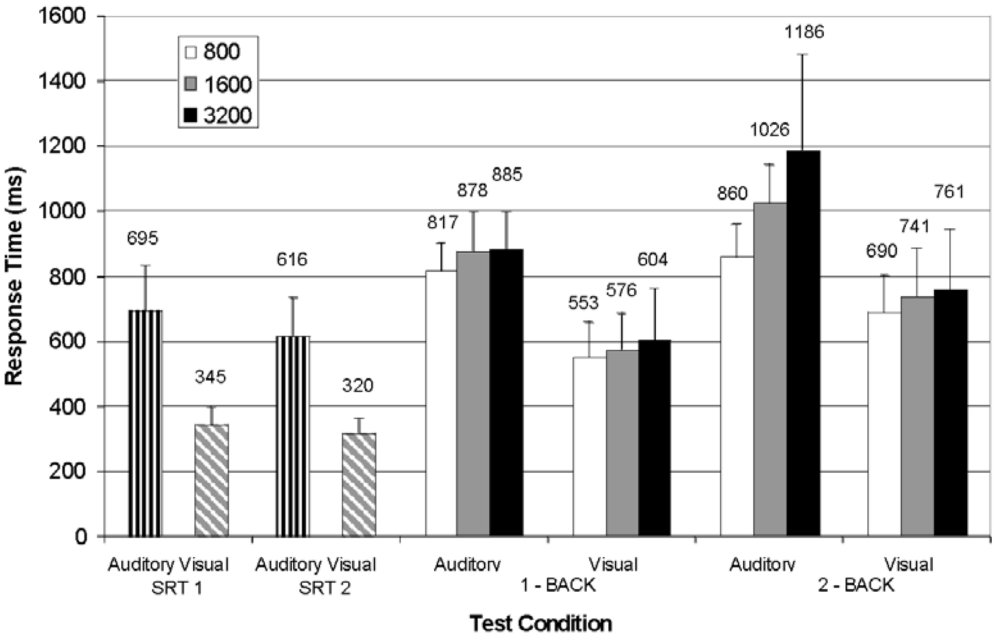
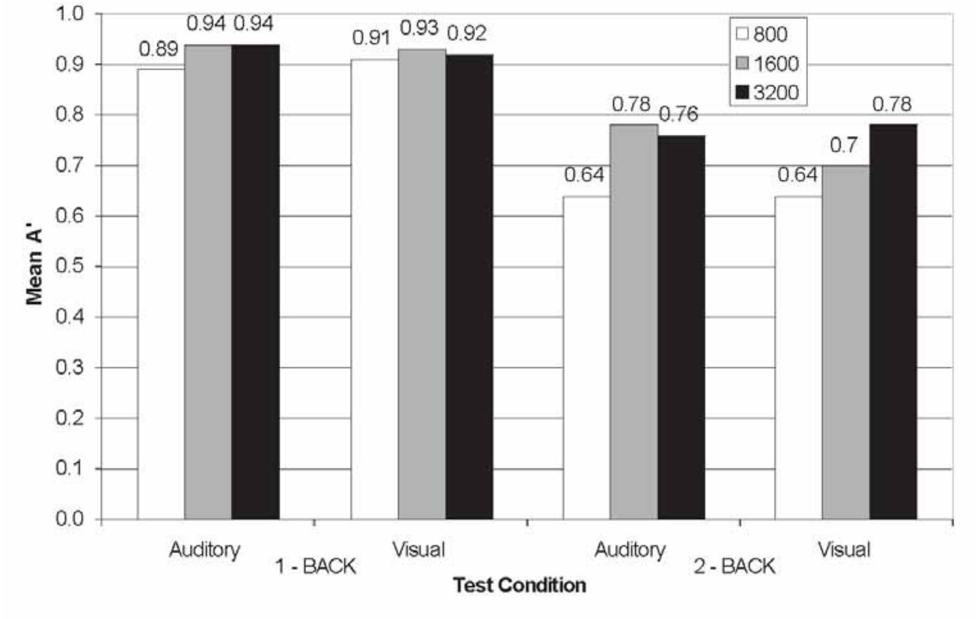


Figure 2. Mean response time (+*SD*) in four tests, two modalities, and three levels of ISI.

Figure 3. Mean sensitivity (*A´*) in *n*-back tasks, two modalities, and three levels of ISI.



response speed. The CWM scores for both groups are displayed in Figure 4. A positive CWM score indicates that 2-back RT was longer than 1-back RT in that condition.

Differences between CWM across the three levels of ISI reached statistical significance at the .05 level, $F(1.46, 66.971) = 13.605$, $p < .000$, $\eta_p^2 = .228$. Pairwise comparisons revealed significant differences between 800 and 1600 ISI levels and between 800 and 3200 ISI levels, but not between 1600 and 3200 ISI levels. There was a significant interaction between modality and ISI, $F(1.66, 76.5) = 4.843$, $p = .015$, $\eta_p^2 = .095$, most likely because of the variability across ISI in the auditory condition. Differences between CWM in auditory and visual modalities did not reach statistical significance, $F(1, 46) = .080$, $p = .779$, $\eta_p^2 = .002$.

The CWM analysis was designed to isolate the effect of working memory, specifically the added effect of updating and recalling two letters previous to the stimuli instead of only one. An intriguing pattern of CWM emerged; although CWM was steady across visual presentations, it ranged from 42 ms during 800 ISI up to 301 ms during 3200 ISI in the auditory conditions.

DISCUSSION

Working Memory Tests

Investigations of working memory in language and other tasks have used a variety of tasks that require both storage and manipulation of information (Waters & Caplan, 2003). The *n*-back task is relatively easy to administer and to record responses via computer program, as in the current study. It is generally accepted that *n*-back tasks requiring more items to be stored in working memory will consume more resources and reduce performance more than those tasks requiring less items (Braver et al., 1997; Jonides et al., 1997; Schumacher et al., 1996). In the current study, as expected, participants' performance on the 2-back task was slower and less accurate than performance on the 1-back task.

Table 2. Pairwise comparisons for mean RT across two *n*-back tasks, two modalities, and three levels of ISI.

	Mean difference	*Standard error*	p
Auditory & visual	285.625*	14.765	<.000
1-back & 2-back	138.715*	13.196	<.000
800 ISI & 1600 ISI	88.880*	8.702	<.000
800 ISI & 3200 ISI	148.635*	15.081	<.000
1600 ISI & 3200 ISI	59.755*	11.446	<.000

Note. Bonferroni adjustments for multiple comparisons.

*significant at the .05 level.

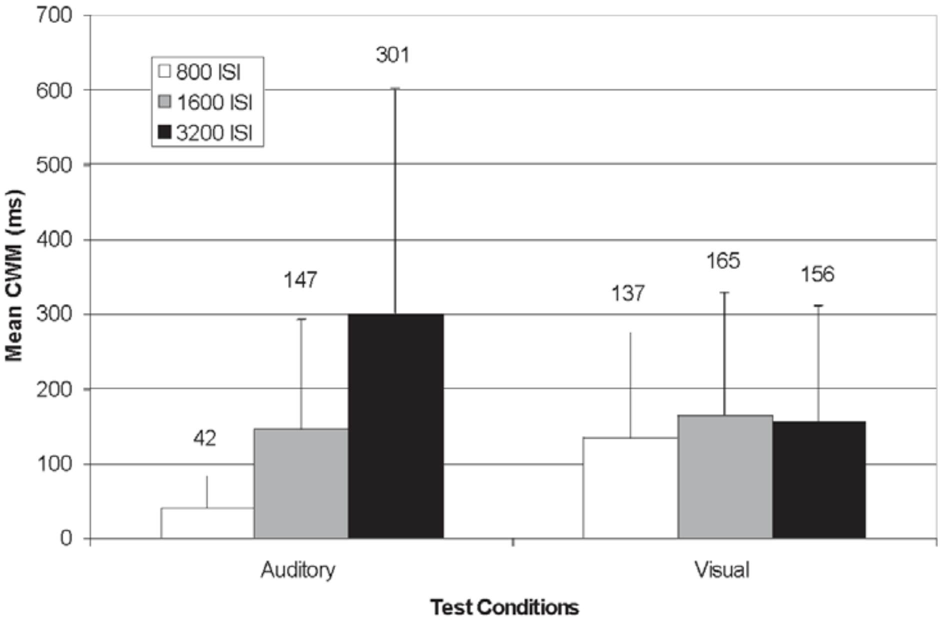
Table 3. Pairwise comparisons for *A´* across two *n*-back tasks, two modalities, and three levels of ISI.

	Mean difference	*Standard error*	p
Auditory & visual	.006	.010	.566
1-back & 2-back	.191*	.012	<.000
800 ISI & 1600 ISI	.046*	.009	<.000
800 ISI & 3200 ISI	.074*	.010	<.000
1600 ISI & 3200 ISI	.027*	.007	.002

Note. Bonferroni adjustments for multiple comparisons.

*significant at the .05 level.

Figure 4. Mean computed working memory (2 back – 1 back) RT (ms) in two modalities and three levels of ISI.



CWM RT is the additional time required to complete the 2-back task compared to the 1-back task. Theoretically, this reflects the time required to process the working memory component without the perception and motoric response portion of the total RT. CWM was not different in the two modalities, but the main effect of ISI was statistically significant. The interaction between modality and ISI was significant, with the source of interaction displayed clearly in Figure 4. In the auditory condition only, RT was slower in the 2-back task by only 42 ms in the 800 ISI condition, but steadily increased to 301 ms slower in the 3200 ISI condition. These two points represent the extremes in the range of CWM RTs. It is difficult to determine why CWM diverged to such an extent. The effect of ISI must have been exaggerated in the 2-back condition. It is possible that participants had more difficulty perceiving the auditory stimuli and took advantage of the longer ISI to spend more time deciding if each stimulus was correct. Therefore, it is possible that this apparently odd result is actually an indication of a strategy that was necessary in the auditory but not visual conditions.

In the current study, CWM was an informative measure contributing another perspective of differences in the *n*-back tasks that were less obvious in the initial analysis. The *n*-back task is a unique measure of working memory because of the linear pattern of complexity and demand created by such a simple modification of adding one number to the set to be stored and manipulated. The stair-step pattern of RT was significant in the current study and was highlighted by the CWM RT pattern (see Figure 4). In the visual conditions, the patterns of modality and ISI in CWM were similar to results from analyses of the total RT and accuracy measures.

Influence of Modality

The influence of stimulus modality in the current study varied depending on the measure of interest. Although measures of accuracy (i.e., hit rate, *A´*) were not significantly affected by the modality of stimuli presented, RT was longer with auditory stimuli. This discrepancy indicates that participants were able to perform all of the tests but required more time during the auditory conditions. Processing of auditory stimuli was somehow more complex, requiring more time, without being too complex, which would cause errors. Because this modality difference was evident in SRT tests as well as *n*-back tests, the difference is more indicative of different perceptual processing times than complexity of working memory. It probably required more time to hear the entire presentation of the aurally presented letter and perceive the auditory stimuli than was required to perceive the visually presented letter.

Schumacher et al. (1996) reported the same findings from several *n*-back tasks and also attributed the discrepancy to encoding differences. Using PET, Schumacher et al. found similar neural patterns of activation in conditions of visual and auditory presentation with the exception of activation in the occipital lobe (i.e., BA18) only during visual presentation and activation in the temporal lobe (BA 21 and 22) only during auditory presentation. Smith et al. (1998) reviewed these findings and attempted to remove the input differences by deriving images of activation in the 3-back condition minus activation in control conditions. A single site in Broca's area remained active after the computed visual image was subtracted from the computed auditory image. The behavioral results from the current computerized assessment concur with the conclusion of Schumacher et al. and Smith et al. (1998): Only the translation of stimuli into phonological representations is modality dependent, and the working memory process is amodal.

The effect of modality was consistent across tasks, indicating that any additional time used for auditory perception was universal and need not invalidate results from studies using auditory stimuli. However, direct comparisons of studies or tests using different modalities of stimuli presentation should heed caution or perhaps be avoided completely. ISI in auditory conditions should allow for additional perceptual processing time. The pragmatic choice when designing a study is to present only visual stimuli, thereby avoiding the risk of participants lacking or varying in auditory acuity or processing speed as well as avoiding the potential to add variability by using different voices (e.g., male or female) or intensity levels across studies. Perhaps this is why most of the CPT studies summarized by Riccio et al. (2001) use visual stimuli. Although a few participants reported difficulty with hearing some letters, which led to a time-consuming outlier analysis and loss of data, none reported difficulty with lowercase or uppercase letters, or the random presentation of lowercase and uppercase. In summary, RT should not be compared across tests using different stimulus modalities, and the visual modality can be used with less concern about sensory variability in participants.

Influence of ISI Duration

Time during ISI may be helpful if the participant uses it to mentally update or repeat previous stimuli subvocally or implement some other sort of strategy, but it also may be detrimental if it is so long as to create a burden on the storage and processing capacities of working memory by having to maintain information for a longer period, perhaps even using some short-term memory (Stuss et al., 2005). Conversely, a short ISI may be helpful by reducing strain on the storage and processing capacities or detrimental by not allowing time and increasing processing demands to update information and inhibit old or irrelevant stimuli quickly (Silverstein et al., 2004). The current study's determination of what qualified as short or long was based on the few studies that have reported ISI, but even the studies including multiple ISIs often do not provide rationale for selection. The minimal concern for ISI influence in the literature implies that the effects of ISI are disregarded, acknowledged without proposed solutions, or not yet recognized.

Significant effects of ISI were found in this study. Participants in this study used more time but also became more accurate with each increase of ISI duration. The significant interaction between *n*-back task and ISI, coupled with visual inspection of Figures 2 and 3, indicates that these effects were particularly evident for the more complex 2-back task. The current study is consistent with previous reports that a short ISI compromises performance findings; longer ISI allowed participants to have the amount of time necessary for processing and responding accurately (Chee et al., 1989; Riccio et al., 2001; Silverstein et al., 2004). It is possible that the short ISI caused participants to be more conscious of their speed and motivated them to make a response decision quickly, before more stimuli were deposited in the filling working memory bank. Perhaps attending to accuracy and balancing speed and accuracy was more manageable in longer ISI conditions when participants did not feel the urgency of preparing for the next stimuli. This hypothesis is difficult to test but is supported by the pattern of results for lengthening ISI and general participant comments such as "It was hard to keep track of [the letters] when they were going so fast" and "It was hard to keep up sometimes."

The findings presented here highlight the importance of using an appropriate ISI, or at least considering the contribution of ISI to a participant's performance, and reporting multiple parameters of performance, such as RT with accuracy. Although a standard ISI has not been established in CPTs, perhaps because an appropriate ISI may vary by task and tested population, researchers sharing results in literature or presentations should report ISI and other parameters for purposes of interpretation and comparison to similar studies as well as for replication.

If the current study is used as a guideline, participants were rushed during the 800 ISI condition and made more errors when responding quickly. Accuracy was very high in the 3200 ISI condition, particularly in the 1-back task. The discrepancy between accuracy in the ISI levels indicates that participants were capable of performing, but a short ISI taxed their capability, resulting in underestimations of their capabilities. The 3200 ISI condition was not long enough to become more detrimental to accuracy than shorter ISI, as might be predicted if 3200 ms was long enough to strain working memory limitations. Performance in the 3200 ISI condition only approached 100% accuracy, or ceiling effects. The nonsignificant difference between 1600 ISI and 3200 ISI in the CWM analysis implies a leveling in the amount of time sufficient for working memory processing in 1- and 2-back tests. The results of this study indicate 3200 ISI to be a reasonable ISI in 1- and 2-back tests, given that participants were able to achieve the highest accuracy in this condition, but not so much higher than 1600 ISI that a continuing linear trend is apparent in the CWM analysis.

Duration of ISI is one way that ISI can influence results. The ISI in this study was presented in blocks and counterbalanced across participants, and was therefore predictable. According to Stuss et al. (2005), when ISI is blocked, RT may increase with longer ISIs because it is difficult to time and maintain peak preparedness over longer ISIs (no quantification of "longer" is provided by Stuss et al.). When Stuss et al. reported a decrease in RT as ISI increased, the ISIs of 4 s to 7 s were random (unblocked) and warning signals prepared the participants for response. A change in RT with changes in the time between a warning signal and the stimulus is known as the foreperiod effect (Los, Knol, & Boers, 2001). Duration of ISI must be considered in context of the entire experiment, including the construct measured, whether ISI is blocked or random, and whether warning signals are present to aid peak preparedness. All of these variables should be included when reporting results to aid in study comparisons and perhaps explain conflicting results. Overall, the conclusions from this study can be summarized as follows:

- Responses in 2-back tasks were later and less accurate than responses in 1-back tasks.
- Modality of stimuli presentation in *n*-back tasks affected RT but not accuracy.
- ISI can influence performance profiles, particularly if the ISI is too short to allow for complete perception and processing of the stimuli.

FUTURE DIRECTIONS

Innovative research using technological advancements is shadowed by basic methodological variations, such as presentation of stimuli, which can confound outcomes. The specific aim of this study was to explore possible confounding parameters of stimuli in computerized *n*-back tasks addressing working memory in healthy people. More generally, this study highlighted the need to carefully select and understand such tools.

This methodological study is preparation for future studies addressing the use of computerized measures for clinical populations. Perhaps more sensitive measures or composite measures of working memory could distinguish the groups before substantial cognitive impairment is detected (Waters & Caplan, 2003). Clinical implications of deteriorating working memory clearly include language and communication impairments as well as adjustments in plan of care. Furthermore, as cognitive–linguistic impairments continue to be recognized and quantified in various clinical populations, more study of appropriate assessment and efficacious treatment is necessary. These tests will need to be modified for computerized response. Designs of future studies of working memory incorporating computerized measurement can be guided by the findings presented here regarding stimulus modality and ISI.

REFERENCES

Ashcroft, M. H. (2002). *Cognition* (3rd ed). Englewood Cliffs, NJ: Prentice Hall.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York: Academic Press.

Barch, D. M., Braver, T. S., Nystrom, L. E., Forman, S. D., Noll, D. C., & Cohen, J. D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia, 35*(10), 1373–1380.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage, 5*, 49–62.

Chee, P., Logan, G., Schachar, R., Lindsay, P., & Wachsmuth, R. (1989). Effects of event rate and display time on sustained attention in hyperactive, normal, and control children. *Journal of Abnormal Child Psychology, 17*(4), 371–391.

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

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General, 121*(3), 275–277.

Engle, R. W., & Conway, A. R. A. (1998). *Working memory and comprehension*. In R. H. Logie & K. J. Gilhooly (Eds.), *Working memory and thinking* (pp. 67–92). East Sussex, UK: Psychology Press.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-Mental State.” A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198.

Jonides, J. (1995). Working memory and thinking. In E. E. Smith & D. N. Osherson (Eds.), *Thinking: An initiation to cognitive science* (Vol. 3, 2nd ed., pp. 215–265). Cambridge, MA: MIT Press.

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., et al. (1997). Verbal working memory load affects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience, 9*(4), 462–475.

Jurica, P., Leitten, C., & Mattis, S. (2001). *The Dementia Rating Scale, 2nd ed. manual*. Odessa, FL: Psychological Assessment Resources.

Kay Elemetrics Corporation. (2002). *Software instruction manual Multi-Speech Model 3700, CSL Models 4100, 4300B, and 4400* (Version 2.5). Lincoln Park, NJ: Author.

Koster, W. G. (Ed.). (1969). Donders, F. C. (1868–1869). On the speed of mental processes. *ActaPsychologica, 30*, 412–431.

Lehman, M. T., & Tompkins, C. A. (1998). Reliability and validity of an auditory working memory measure: Data from elderly and right-hemisphere damaged adults. *Aphasiology, 12*(7/8), 771–785.

Los, S. A., Knol, D. L., & Boers, R. M. (2001). The foreperiod effect revisited: Conditioning as a basis for nonspecific preparation. *Acta Psychologica, 106*, 121–145.

Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General, 130*(2), 199–207.

MacDonald, C. J., & Meck, W. H. (2004). Systems-level integration of interval timing and reaction time. *Neuroscience and Biobehavioral Reviews, 28,* 747–769.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and “nonparametric” indexes. *Psychological Bulletin, 107*(3), 401–413.

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin, 98*(1), 185–199.

May, C. P., Zacks, R. T., Hasher, L., & Malthaup, K. (1999). Inhibition in the processing of garden-path sentences. *Psychology and Aging, 14*(2), 304–313.

Nystrom, L. E., Braver, T. S., Sabb, F. W., Delgado, M. R., Noll, D. C., & Cohen, J. D. (2000). Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. *NeuroImage, 11*, 424–446.

Okazaki, S., Hosokawa, M., Kawakubo, Y., Ozaki, H., Maekawa, H., & Futakam, S. (2004). Developmental change of neuro-cognitive motor behavior in a continuous performance test with different interstimulus intervals. *Clinical Neurophysiology, 115*(5), 1104–1113.

Riccio, C. A., Reynolds, C. R., & Lowe, P. A. (2001). *Clinical applications of continuous performance tests*. New York: John Wiley & Sons.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-prime (Version 1.0) [Computer software]. Pittsburgh, PA: Psychology Software Tools.

Schumacher, E. H., Lauber, E., Awh, E., Jonides, J., Smith, E. E., & Loeppe, R. A. (1996). PET evidence for an amodal verbal working memory system. *NeuroImage, 3,* 79–88.

Silverstein, M. L., Weinstein, M., & Turnbull, A. (2004). Nonpatient CPT performance varying target frequency and interstimulus interval on five response measures. *Archives of Clinical Neuropsychology, 19*, 1017–1025.

Smith, E. E, Jonides, J., Marshuetz, C., & Koeppe, R. A. (1998). Components of verbal working memory: Evidence from neuroimaging. *Proceedings of the National Academy of Sciences, 95,* 876–882.

Sternberg, S. (2001). Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychologica, 106*, 147–246.

Sternberg, S. (2003). Process decomposition from double dissociation of subprocesses. *Cortex, 39,* 180–182.

Stuss, D. T., Alexander, M. P., Shallice, T., Picton, T. W., Binns, M. A., Macdonald, R., et al. (2005). Multiple frontal systems controlling response speed. *Neuropsychologia, 43*, 396–417.

Towse, J. N., Hitch, G. J., & Hutton, U. (2000). On the interpretation of working memory span in adults. *Memory and Cognition, 28*(3), 341–348.

Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavioral Research Methods, Instruments, & Computers, 35*(4), 550–564.

Wechsler, D. (1997). *Weschler Adult Intelligence Scale—III*. San Antonio, TX: The Physhological Corporation.

Wei, X., Yoo, Dickey, C., Zou, K. H., Guttmann, C. R. G., & Panych, L. P. (2004). Functional MRI of auditory verbal working memory: Long-term reproducibility analysis. *Neuro-Image, 21,* 1000–1008.

Contact author: Adrienne B. Hancock, Speech and Hearing Sciences Department, 2115 G Street NW, Washington, DC 20052. E-mail: Hancock@gwu.edu.