Published in final edited form as:
J Clin Exp Neuropsychol. 2010 March ; 32(3): 315–323. doi:10.1080/13803390903032529.

Can We Improve the Clinical Assessment of Working Memory? An Evaluation of the WAIS-III Using a Working Memory Criterion Construct

Benjamin Hill, **Emily Elliott**, **Jill Shelton**, **Russell Pella**, **Judith O'Jile**, and **William Gouvier**

Abstract

Working memory is the cognitive ability to hold a discrete amount of information in mind in an accessible state for utilization in mental tasks. This cognitive ability is impaired in many clinical populations typically assessed by clinical neuropsychologists. Recently, there have been a number of theoretical shifts in the way that working memory is conceptualized and assessed in the experimental literature. This study sought to determine to what extent the WAIS-III working memory index (WMI) measures the construct studied in the cognitive working memory literature, whether an improved WMI could be derived from the subtests that comprise the WAIS-III, and what percent of variance in individual WAIS-III subtests is explained by working memory. It was hypothesized that subtests beyond those currently used to form the WAIS-III WMI would be able to account for a greater percentage of variance in a working memory criterion construct than the current WMI. Multiple regression analyses (n = 180) revealed that the best predictor model of subtests for assessing working memory was composed of the Digit Span, Letter-Number Sequencing, Matrix Reasoning, and Vocabulary. The Arithmetic subtest was not a significant contributor to the model. These results are discussed in the context of how they relate to Unsworth and Engle's (2006, 2007) new conceptualization of working memory mechanisms.

Keywords

Neuropsychological Assessment; Working Memory; WAIS-III; Dual-Component Model

Working memory refers to the cognitive system that stores information in an accessible state for utilization in complex mental tasks. This system also allows for updating and manipulation of relevant information. It is composed of both passive storage and dynamic control processes in order to hold information in an active form (Baddeley, 2001). Working memory is impaired in a wide variety of neuropsychiatric conditions, including dementia (Collette, Van der Linden, & Salmon, 1999), attention-deficit hyperactivity disorder (ADHD; Pasini, Paloscia, Alessandrelli, Porfirio, & Curatolo, 2007), and schizophrenia (Fleming, Goldberg, Gold, & Weinberger, 1995; Goldman-Rakic, 1994) making its accurate assessment an important clinical issue.

Baddeley originated our current view of working memory and his conceptualization (Baddeley & Hitch, 1974) still dominates the field. Baddeley's model proposed a central executive component to manipulate data via controlled attention processes. It also contained two supportive systems that serve the central executive by being able to either passively hold or actively rehearse information. These two systems are termed the phonological loop and

---

Contact: Benjamin Hill, Clinical Neuropsychology Fellow, Alpert Medical School of Brown University - Department of Psychiatry and Human Behavior, Providence, Rhode Island 02914, United Statesbhill@lifespan.org.

visuospatial sketchpad, and hold verbal information and visuospatial information respectively. More recently, Baddeley amended the model by including an interface, the episodic buffer, which allows the working memory system to communicate with long-term memory (for a review see Baddeley, 2001).

Since Baddeley's original model was proposed, several new conceptualizations of working memory have emerged (Ericsson & Kintsch, 1995; Engle & Oransky, 1999; Cowan, 2001, 2005; Oberauer, 2002). Perhaps the most intriguing of these from a clinical standpoint is Unsworth and Engle's (2006, 2007) recent dual-component model of working memory ability. In Engle's original framework (Engle & Oransky, 1999), working memory was seen as being composed of two modules: immediate memory and controlled attention (sometimes called executive attention). However, Unsworth and Engle (2007) have recently proposed a heavily modified version of Engle's original model. Working memory is now conceptualized as an interaction between two modules: 1) a primary memory storage buffer that maintains up to four chunks of information for a limited amount of time (though it can focus on fewer chunks) and 2) a cue-dependent search system that retrieves data from longer-term memory, which they term secondary memory. They have postulated that all information is initially stored in primary memory but, whenever this system reaches capacity, information is automatically moved to secondary memory.

According to this model, individual differences in working memory ability are directly related to the efficiency with which a person can search the secondary memory system via cues, the most prominent of which are temporal-contextual cues. Unsworth and Engle (2005, 2007) still speculate that attentional control mechanisms play a role in this system by increasing search efficiency in secondary memory. They have also asserted that working memory and fluid intelligence are strongly related as both constructs rely on both executive attention and the capability to efficiently search secondary memory (Unsworth & Engle, 2006). Others have also proposed that working memory and fluid intelligence may be functionally related (Kyllonen & Christal, 1990; Engle, Kane, & Tuholski, 1999).

Clinical Assessment

Currently, the Wechsler Adult Intelligence Scale, 3rd editon (WAIS-III) is the most commonly used instrument in the armamentarium of clinical neuropsychologists (Camara et al., 2000; Rabin, Barr, & Burton 2005). Therefore, the Working Memory Index (WMI) from that battery is commonly used to assess working memory ability. The WAIS-III WMI is derived from the performance on the Digit Span, Arithmetic, and Letter-Number Sequencing subtests. These subtests will now be discussed further.

The Digit Span subtest is composed of both a forward and a backward recitation condition. On the digits forward part of the subtest, the individual is verbally presented with a string of numbers and asked to repeat back the numbers in order immediately after stimuli presentation. In the digits backward condition, the individual is instructed to repeat back the presented string of numbers in reverse order. The Digit Span score combines the total number of digit strings correctly repeated in both conditions.

Although Digit Span is conceptualized as a test of working memory on the WAIS-III, it seems to have a heavy immediate memory component inherent in its design, particularly the Digit Span forward condition. Also, Reynolds (1997) noted that the forward and backward components of digit span tasks loaded on separate factors raising the possibility that the task measures more than just working memory ability. Additionally, this subtest appears to have low face validity as a measure of working memory due to its lack of task complexity. However, as previously discussed, recent research has supported the idea that both immediate memory and working memory may share a number of cognitive processes including those required to perform simple span tasks (Unsworth & Engle, 2006).

The Arithmetic subtest is also used to calculate the WAIS-III WMI. On this measure, the individual is verbally presented with arithmetic word problems that increase in difficulty. While this task does appear to manipulate working memory load as part of the assessment procedure, and it is not an assessment of mathematic ability per se, it is likely influenced by mathematical abilities to some degree making it a less than ideal paradigm for assessing working memory performance.

In contrast to the two previously discussed subtests, the Letter-Number Sequencing subtest has high face validity as a working memory task as working memory is generally conceptualized. In this task, the individual is presented with a series of numbers and letters in random order. They are then instructed to repeat back the numbers in order first followed by the letters in alphabetical order. Unlike the Digit Span subtest, this subtest appears to require more than just immediate memory and, unlike the Arithmetic subtest, there is no minimum academic skill prerequisite other than knowing the numbers 1 through 9 and having a functional knowledge of the alphabet.

In sum, the WMI derived from the WAIS-III may not reflect the best approximation of working memory ability as conceptualized by current theoretical models. This is mainly due to the possibility that there may be a high amount of variance associated with immediate memory or non-working memory factors in these measures. Based on earlier conceptualizations of working memory (Baddeley, 1974; Engle & Oransky, 1997), this situation could potentially be significantly rectified by removing subtests that may be measuring primarily passive immediate memory. Conversely, newer theories of working memory (Unsworth & Engle, 2006) predict that simple span tasks that require the individual to hold over four chunks of information will be functionally equivalent to other working memory tasks. Another potential weakness of the subtests that comprise the WAIS-III WMI is that their methodology departs substantially from the working memory tasks commonly utilized in the experimental literature by cognitive psychologists.

Experimental Assessment

In experimental psychology and cognitive neuroscience, working memory is typically assessed via tasks that combine both storage and processing tasks in order to fully assess working memory as it is theoretically conceptualized. Such tasks are known as complex span tasks. Performance on these complex span tasks is typically both reliable and highly correlated across tasks (Klein & Fiss, 1999; Daneman & Merickle, 1996; Conway et al., 2005). One of the first developed complex span tasks was the sentence span task (Daneman & Carpenter, 1980). In this task, the participant must indicate that they comprehend a presented sentence and retain the last word of the sentence. Sentence comprehension represents the secondary ongoing attentional component of the task while recall of the last word in each sentence represents the storage function of the task. Working memory capacity is defined as the number of sentences that are both correctly comprehended and their last word correctly recalled in order. If the sentence is read by the participant, the task is called the reading span task and if the sentence is presented orally to the participant, the task is called the listening span task (l-span).

Turner and Engle (1989) built on this paradigm, developing the operation span task (o-span). In the o-span task, an arithmetic problem is presented followed by a word to be remembered. The number of words recalled in the correct order represents the working memory span. Similar to the sentence span task and unlike the clinical Wechsler subtests, there is a secondary ongoing attentional task in the solving of arithmetic problems. Additionally, varying arithmetic difficulty of the o-span task to equalize performance across participants has been shown to not affect the relationship between o-span and reading span performance indicating that a common process is being assessed across tasks (Conway & Engle, 1996).

The *n*-back task, also known as the lag task, has also been commonly employed in the experimental literature as a measure of working memory ability, to the point that it has been referred to as the gold standard working memory assessment technique in cognitive neuroscience (Kane & Engle, 2002). In the *n*-back task, individuals are presented with stimulus sequences and required to recall a stimulus presented a specified number back in the sequence (*n* represents how far back in the sequence the person is to go). For example, at an *n* of one, the correct response would be whatever was presented immediately prior to the last stimulus. At an *n* of two, the correct response is what was presented two prior to the last stimulus. This task has high face validity as a measure of working memory due to the attentional requirement to continuously update the stimuli being held on-line.

After reviewing the methods used to measure working memory capacity in clinical versus experimental settings, it becomes clear that there are significant differences between these two paradigms. In particular, the clinical subtests appear to lack the simultaneous processing component that taxes attentional resources and is a hallmark of the experimental cognitive tasks. Additionally, recent research has demonstrated that, with the exception of Letter-Number Sequencing, the other WAIS-III subtests may not be measuring the same working memory construct that is studied in the cognitive literature (Shelton, Elliott, Hill, Calamia, & Gouvier, 2009). Cognitive psychologists defined the term “working memory” and have studied the reliability and validity of their tasks in depth to ensure that they are accurately measuring the working memory construct (Conway et al., 2005). As the cognitive literature is commonly translated to the clinical situation, it is important to both demonstrate that clinicians are assessing an equivalent working memory construct and to establish the specificity with which we measure working memory.

This study seeks to answer three specific questions: 1) to what degree are clinicians currently assessing the working memory construct that is studied in the cognitive literature, 2) can an improved WMI be derived from the subtests that comprise the WAIS-III, and 3) what percent of variance does working memory account for in individual WAIS-III subtests? Currently, only a small number of subtests from these two instruments are used to assess working memory despite the fact that many other Wechsler subtests measure different aspects of attention and concentration abilities (e.g., Matrix Reasoning). The abilities assessed by other subtests not included in the WAIS-III WMI are likely functionally related to executive attention and retrieval from secondary memory aspects of working memory that were previously discussed. It is hypothesized that a model composed of subtests from the WAIS-III beyond those currently used to form the WMI, particularly subtests related to fluid intelligence, will be able to account for a greater percent of variance in a working memory criterion construct than the current WAIS-III WMI.

Methods

Participants

Participants were compensated by receiving extra credit in a course offered in the Department of Psychology of a state university. No participants were excluded based on gender, race, ethnicity, or academic status. However, the following factors did exclude an individual from inclusion in this study: a visual and/or hearing impairment at the time of testing, a relevant psychiatric diagnosis that would result in cognitive impairment, identifying English as their secondary language, or failing to attend both experimental sessions. A total of 188 participants met inclusion criteria and agreed to participate in this study.

Of these 188 individuals, the data from eight participants were excluded to performing at a less than 80% correct responding rate (more than 15 errors) on the mathematics section of the o-span task used in this study. Similar response criteria rates have been implemented in other studies using this task (Unsworth & Engle, 2005, 2006) as such performance may indicate a less than acceptable level of engagement in the testing procedure; this likely reflects focusing on rehearsal strategies during testing at the expense of the simultaneous processing component. Enforcing this criterion left a sample of 180 participants (mean age in years 20.6, SD 3.6; mean WAIS-III full scale IQ 109.9, SD 11.4; 72.8% female, 27.2% male). While there is a significantly higher proportion of females than males in this sample ($\chi^2 = 40.62$, $p < .001$), the literature states that there is no significant gender difference in performance on any of the measures used in this study (Psychological Corporation, 1997; Robert & Savoie, 2006). Additionally, 7 participants had missing data that was imputed. The process used to impute these missing values will be explained further in the Results section.

Measures

The following measures were utilized in this study.

WAIS-III—The following subtests of the WAIS-III (Wechsler, 1997a) were administered to all participants according to the protocol described in the test manual: Picture Completion, Vocabulary, Digit Symbol Coding, Similarities, Block Design, Arithmetic, Matrix Reasoning, Digit Span, Information, Picture Arrangement, Comprehension, Symbol Search, and Letter-Number Sequencing. While index scores such as Full Scale IQ and the WMI were standardized in respect to the subject's age according to standardized normative data found in the test manual. Raw score data was utilized for analyses of individual WAIS-III subtests.

Automatic Operation Span Task—For this study, a modification of the o-span task was utilized, referred to as the automatic o-span task (ao-span; Unsworth, Heitz, Schrock, & Engle, 2005). In the ao-span, a letter and a simple, two-step mathematics problem were simultaneously presented to participants on a computer screen. They were instructed to both retain the presented letter and correctly respond in a true or false format to a number presented as the answer to the mathematic problem. The correct mathematical answer was presented half the time randomly. A list of 3–7 letters (3 trials at each length) was presented randomly during the task and participants were instructed to select the letters that they had previously been asked to retain in the order of presentation. Participants were reminded periodically throughout the task to maintain a high level of correct responding to the mathematics problem. This ensured activation of the secondary processing component and decreased the probability that they were simply rehearsing the list of presented letters. The ao-span score that was used in this study is an index that reflects the number of letters recalled on errorless trials with more weight given to longer and more difficult trials.

Listening Span Task—The l-span task (Cowan et al., 2005) utilized in this study was based on the paradigm utilized by Kail and Hall (1999). This was a modification of the reading span task (Daneman & Carpenter, 1980) where the stimuli were presented via headphones rather than being read by participants. In the l-span, participants were presented with an auditory sentence and had to then press a button characterizing the sentence's content as either true or false. They were also instructed to retain the last word of each presented sentence. Randomly between two to nine sentence presentations, the subject was instructed to type the last word of each sentence in sequence. The subject completed three trials at each span length (two to nine sentences), resulting in a trial block. The task was discontinued when three incorrect responses were given at any trial block. The l-span score was calculated by summing the number of words recalled on errorless trials.

Modified Lag Task—A version of the *n*-back task, known as the modified lag task (MLT), was used (Dobbs & Rule, 1989; Shelton, Metzger, & Elliott, 2007). The MLT has an advantage over other forms of the *n*-back task in that it has demonstrated convergent validity with other measures of working memory (Shelton et al., 2007). In the task, subjects were shown lists of words one word at a time. List length was either four or six words. All words used in the MLT were matched for familiarity and frequency of usage. Immediately after viewing the list, participants were asked to type either the last word from each list, the word that was one back in each list, two back in each list, or three back in each list. Each participant was given five trials at each of the four *n*-back conditions (0-back, 1-back, 2-back, and 3-back) for both the list length of four and the list length of six with each participant completing 40 trials. The MLT score that was used is an index that reflects performance on errorless trials with more weight given to more difficult *n*-back conditions (e.g. 0-back performance × 1, 1-back performance × 2, 2-back performance × 3, and 3-back performance × 4).

Procedure

Data collection for this study was approved by an Institutional Review Board. All data were collected in two sessions that took place approximately one week apart. This was necessary due to the amount of time required for data collection. All experimental working memory tasks were presented via desktop computer.

Results

Treatment of Missing Data

For missing data, maximum likelihood estimation with an expectation maximization algorithm (Jamshidian & Bentler, 1999) was used for imputation using the EQS 6.1 program (Bentler, 1995). Expectation maximization methods are superior to other imputation methods, such as regression, because they do not overfit the final solution and produce more realistic estimates of variance (Ullman, 2001). Also, maximum likelihood estimation has more lenient statistical assumptions. A total of seven cases contained missing data with the majority of cases having only one variable missing.

Preliminary Data Analysis

Descriptive statistics are presented for all of the measures that were given in this study (see Table 1 and 2). All of the measures were normally distributed and had reasonable skewness and kurtosis coeffecients. The data were assessed for outliers by converting all utilized variables to z-scores. For the data converted to z-scores, scores in excess of 3.29 in either direction were considered potential outliers (Tabachnick & Fidell, 2001). Six such cases were discovered and their scores were replaced by the next closest score for that variable that was within 3.29 SD. Additionally, the WAIS-III data were checked for reliability by assessing the scoring forms directly and rescoring all data; high scoring reliability in excess of $r = .90$ was found for all variables. Such a reliability check was not necessary for the ao-span, l-span, and MLT data as these tasks were computer administered.

Working Memory Criterion Construct Score

The cognitive experimental measures were utilized as a working memory criterion construct for this study. The ao-span, l-span and MLT data were subjected to an exploratory factor analysis (EFA). Principal axis factoring was used to assess the underlying construct. Such a data reduction approach reduces both method variance and error variance found in each of the three individual experimental cognitive measures and provides a single relatively “pure” measure of their common variance. A single factor emerged (see Table 3) that explained 49.75% of shared variance in the three variables. All statistical assumption were met including the Kaiser-Meyer-Olkin measure of sampling adequacy being sufficient (.67) and Bartlett's test of sphericity being significant [χ^2 (3, N=180) = 121.47, $p<.001$]. This factor was saved as a regression derived factor score for each participant using SPSS 15.0 (2006). This factor score is utilized in further analyses as the working memory criterion construct (WMCC) score.

Multiple Regression

To address the study hypothesis that additional WAIS-III subtests beyond those currently used to form the WMI can be used to better measure working memory ability, a stepwise multiple regression was conducted to determine the best linear combination of subtests from the WAIS-III for predicting the WMCC score. This was hypothesized to included measures of fluid intelligence. Multiple regression using backward elimination, which initially includes all predictor variables before eliminating the least significant predictor, refitting the model, and beginning the process again until only the best model of predictors remains, was used to determine which WAIS-III subtests were the best predictors of working memory ability. Additionally, to reduce potential type I errors, the p value model entry criterion was set at .025 and the model removal p value criterion was set at .05. A separate regression was used to determine the amount of variance the WAIS-III WMI accounted for the in the WMCC.

WAIS-III Predictor Model—All of the subtests from the WAIS-III were entered into a backward stepwise multiple regression as predictors of the WMCC score. The following combination of subtests significantly predicted [$F(4, 175) = 33.26$, $p<.001$] WMCC score: Letter-Number Sequencing, Matrix Reasoning, Digit Span, and Vocabulary. This model accounted for 43% of the variance in WMCC score (adjusted $R^2 = .42$). The standardized beta weights (see Table 4) suggested that Digit Span and Letter-Number Sequencing contributed the most to predicting the WMCC score. This compares to the WAIS-III WMI score itself accounting for 39% of the variance (adjusted $R^2 = .38$) in the WMCC [$F(1, 178) = 111.44$, $p<.001$].

In summary, the clinical measures do appear to be measuring a working memory construct similar to what is studied in the cognitive working memory literature. Also, the hypothesis of this study that additional subtests, particularly those related to fluid intelligence, could be used to more precisely measure working memory was supported by these results but the amount of additional variance accounted for in working memory was modest. Additional subtests from the Wechsler batteries beyond those currently used to derive the WMI were shown to account for only 4% percent more of the variance in the working memory criterion. Also, Matrix Reasoning, considered a strong measure of fluid intelligence on the WAIS-III (Tulsky, Saklofske, & Zhu, 2003), was found to be a significant predictor of working memory ability. It was also notable that the Arithmetic subtest currently used in the WAIS-III WMI was not included in the best predictor model. .

To determine whether the modest increase in variance accounted for in working memory ability provided by the best fit model regression equation (see Table 4) would be clinically significant, we examined the relationship between the current WAIS-III WMI score and the score derived for each individual from the previously discussed best fit model regression equation. The best fit model regression equation score was standardized by converting it to a z-score. This score correlated with the WMI score $r(178) = .91$, $p<.001$. This would seem to indicate that the two measures are not very different. However, when only those individuals 1 standard deviation beyond the mean of the standardized regression equation score were examined, this relationship dropped to $r(33) = .71$, $p<.001$ for those more than 1 standard deviation above the mean and to $r(31) = .47$, $p=.006$ for those less than 1 standard deviation below the mean. The WAIS-III WMI categorized 26% of the sample as falling at least 1 standard deviation on either side of the mean while the standardized best fit model regression equation score categorized 38% of the sample similarly. This result potentially means that the sensitivity of the two scores differs in those with more extreme scores, exactly the type of individuals clinicians are likely to assess. Overall, the standardized regression equation score disagreed with the WMI score on categorizing an individual as being either greater than 1 standard deviation above the mean or less than 1 standardization below the mean 26% of the time.

A final goal of this study was to determine the amount of variance accounted for by working memory in the subtests that make up the WAIS-III. Separate linear regressions were run for each individual WAIS-III subtest with the WMCC as the predicted variable. These results are presented in Table 5. Working memory accounted for at least 10% of the variance on 7 of the WAIS-III subtests and 27% of the variance in full scale IQ underlying the clinical importance of this variable.

Discussion

Currently, the third edition of the WAIS uses the Arithmetic, Digit Span, and Letter-Number Sequencing subtests to measure working memory ability. The extent to which these subtests validly index working memory has recently been brought into question (Shelton et al., 2009). This study sought to answer three specific questions. First, we wished to determine the degree to which common clinical measures of working memory are indexing the same working memory construct that is evaluated in the cognitive literature. Related to this, the results of this study supported that the WAIS-III WMI and Wechsler subtests such as Digit Span and Letter-Number Sequencing are accounting for a significant amount of variance in a criterion construct composed of three measures of working memory ability commonly used in the cognitive literature. These results help support the premise that cognitive research on working memory can be translated to the clinical domain in the interpretation of the Wechsler subtests utilized in this study.

Second, this study wanted to ascertain whether additional subtests from the WAIS-III that are not included in the calculation of the WMI could be used to improve working memory assessment. As many of the Wechsler subtests assess different facets of attention and concentration, it was hypothesized that other subtests outside of those used to form the WMIs would account for variance in working memory beyond the current WMI subtests. Theoretically, such additional subtests may be able to account for variance in working memory that is attributable to executive attention functions and controlled retrieval of items from secondary memory (Unsworth & Engle, 2006, 2007). It was specifically hypothesized that subtests that measure aspects of fluid intelligence would predict working memory ability based on theoretical considerations of shared cognitive functions (Kyllonen & Christal, 1990; Engle, Kane, & Tuholski, 1999; Unsworth & Engle 2005, 2007). The results of this study modestly supported this hypothesis and will now be discussed.

A multiple regression that focused on the WAIS-III found that the best combination of subtests for assessing working memory were Digit Span, Letter-Number Sequencing, Matrix Reasoning, and Vocabulary. This combination of subtests accounted for 4% more variance in working memory than the WAIS-III WMI. Digit Span and Letter-Number Sequencing were by far the best indicators of working memory ability among the WAIS-III subtests, accounting for 33% and 28% of the variance in the working memory criterion construct respectively. While it was expected that Digit Span and Letter-Number Sequencing would be the strongest predictor variables, it was not anticipated that Digit Span would account for the most variance explained (see Table 5) as Letter-Number Sequencing has more demonstrated validity as a working memory measure (Shelton et al., 2009). However, the strength of Digit Span as a predictor supports previous findings that simple span tasks are roughly equivalent to much more complex working memory tasks (Unsworth & Engle, 2006) as retaining anything more than four chunks of information is believed to require retrieval from secondary memory similar to complex working memory tasks.

More surprising was that Arithmetic was excluded from the model as it is currently included in the WAIS-III WMI and a similar version is used on the new WAIS-IV to measure working memory. The current study also found that Arithmetic only explained 14% of the variance in the working memory criterion construct and its exclusion from the best predictor model suggests that this is not unique variance. As factor analytic studies have found that Arithmetic tends to load across WAIS-III factors and has high intercorrelations with other Wechsler subtests (Tulsky et al., 2003), it is possible Arithmetic is more an index of *g* than a specific working memory measure and may only weakly assess the working memory construct. The current results raise questions concerning its utility as a working memory measure.

The inclusion of Matrix Reasoning is intriguing as it fits well with Engle’s theory that specific aspects of working memory are shared with fluid intelligence (Engle, Tuholski, Laughlin, & Conway, 1999; Kane & Engle, 2002; Unsworth & Engle, 2005, 2006). The Matrix Reasoning subtest from the WAIS-III is generally considered a strong index of fluid intelligence (Tulsky et al., 2003). Table 5 illustrates that it is accounting for 10% of the variance in working memory and its inclusion in the best fit model of WAIS-III subtests predicting working memory ability suggests that some measure of this is unique variance beyond that accounted for by Digit Span or Letter-Number Sequencing. Theoretically, it can be deduced that Matrix Reasoning is possibly accounting for either the controlled attention component of working memory (Engle et al., 1999; Unsworth & Engle, 2005) or retrieval from secondary memory processes (Unsworth & Engle, 2006). This result suggests that future versions of the WAIS should include more measures of fluid intelligence and derive ways to include such subtests in future WMIs. While the WAIS-IV includes more measures of fluid intelligence, these are not utilized in any way to measure working memory ability (Wechsler, 2008). Additionally, Arithmetic is retained in the WAIS-IV (though with new test items) as a component of the WMI while Letter-Number Sequencing is now an optional subtest that is not used to calculate the WAIS-IV WMI. The current results, as well as previous research (Shelton et al., 2009) suggest that this may be an unfortunate occurrence as the current Arithmetic subtest is a poor predictor of working memory ability and Letter-Number Sequencing is a superior predictor of this construct, at least as it is measured in the cognitive literature.

The inclusion of the Vocabulary subtest in the WAIS-III best predictor model of working memory ability makes the least intuitive sense of all the predictor variables. However, it has been reported that tests that assess vocabulary abilities are highly related to an updating function in working memory (Friedman, Miyake, Corley, Young, DeFries, & Hewitt, 2006). Additionally, this subtest is generally considered the best index of *g* among the Wechsler subtests (Psychological Corporation, 1997; Tulsky et al., 2003). A recent large-scale lesion mapping study also found that performance on WAIS indexes of vocabulary comprehension and working memory largely share a common cortical substrate in the left frontal area (Gläscher et al., 2009). Therefore, the inclusion of Vocabulary as a significant predictor of working memory ability may be due to a number of factors. . Its inclusion in the model does suggest that it is accounting for some unique variance in working memory not accounted for by Digit Span, Letter-Number Sequencing, and Matrix Reasoning. However, it is notable that additional analyses found that a model composed of Digit Span, Letter-Number Sequencing, Matrix Reasoning, and Block Design accounted for an identical amount of variance as the model including Vocabulary.

To help determine whether the new derived working memory regression equation score presented in this study differed from the current WMI score provided by the WAIS-III, the relationship between the two scores was examined. While they were highly correlated overall, this relationship changed dramatically when individuals more than 1 standard deviation away from the mean were examined. The effect was particularly pronounced in those with lower working memory scores. This is particularly relevant as 1 standard deviation below the mean is the typical cut-score for mild impairment in neuropsychological assessment. From a classification standpoint, this study found that the new derived working memory score disagreed with the WMI score on whether an individual was either greater than 1 standard deviation above the mean or less than 1 standard deviation below the mean 26% of the time. Other results suggest that the derived working memory regression equation score was more sensitive than the WAIS-III WMI as it categorized a higher percentage of individuals as falling at least 1 standard deviation beyond the mean. Overall, these results suggest that the additional variance accounted for in working memory by the best fit model presented in this study may have clinical utility in indentifying individuals with very low or very high working ability. In particular, it suggests that clinicians could improve their ability to identifying those with impairment in working memory, by using the currently proposed combination of tests or other methodologies.

The third goal of this study was to determine the amount of variance working memory ability accounts for in each of the individual WAIS-III subtests. Working memory was found to account for at least 10% of the variance in seven of the subtests that make up the WAIS-III. It accounted for the most variance in Digit Span and Letter-Number Sequencing, 33% and 28% respectively. A surprising finding was that the working memory criterion used in this study accounted for 27% of the variance in WAIS-III full scale IQ. Overall, these results illustrate how important the accurate assessment of the working memory construct is clinically.

In summary, common subtests utilized by clinical neuropsychologists to measure working memory ability do appear to be indexing the same construct that is being evaluated in the cognitive working memory literature. Also, the clinical assessment of working memory using the WAIS-III may be modestly improved by including additional subtests that are not currently included in the WMI, specifically Matrix Reasoning and Vocabulary. However, even with the addition of these subtests, more than half the variance in working memory was not accounted for. This study did demonstrate that the clinical assessment of cognitive constructs can be improved. Additionally, the regression logarithm presented here provides an alternative way to measure working memory ability using the WAIS-III. Also, working memory appears to account for a significant amount of variance in a variety of subtests commonly used by clinicians and accounted for over a quarter of the variance in full scale IQ.

There are, however, a number of limitations to this study which mainly focus on the sample that was utilized. This research utilized a young adult sample that had above-average full scale IQ as a group (see Table 1). Therefore, the current findings may not generalize to individuals with lower levels of intellectual functioning. Additionally, the age of this sample raises questions about the extent to which these results would extend to older individuals, particularly geriatric populations. Also, the majority of participants demonstrated average to above average working memory functioning. As such, the findings of this study may not apply to individuals with lower levels of working memory ability (Conway & Engle, 1994). Finally, the relatively high level of education of this sample may complicate the generalizability of our attained results given that education is known to be an influential variable in neuropsychological performance (Kaufman, Reynolds, & McLean, 1989). However, it was important to use a sample of college participants as the starting point of this research as a large portion of research in the working memory literature has utilized this segment of the population. Future studies should be conducted with clinical populations.

Still, there are several areas of research that are suggested by the results of the present study. The most obvious is that the field of clinical neuropsychology could potentially benefit from adapting the current experimental cognitive methodologies that are used to measure working memory to make it more applicable to the clinical environment. This would provide clinicians a different and potentially superior assessment paradigm to study pathological working memory impairment in clinical populations. Such an endeavor would also require the collection of normative data for these assessment methods. Related to this, the ability to parse out the controlled attention and secondary memory retrieval processes in working memory would yield another cognitive modality for neuropsychologists to investigate that has not previously been accessed. Future research should continue to address these important issues in the clinical assessment of cognitive functioning.

References

Baddeley, AD. Working memory. Oxford, England: Clarendon Press; 1986.

Baddeley AD. Is working memory still working? American Psychologist 2001;56:849–864. [PubMed: 11785151]

Baddeley, AD.; Hitch, GJ. Working memory. In: Bower, G., editor. Recent advances in learning and motivation. Vol. VIII. New York: Academic Press; 1974.

Barrouillet P, Bernardin S, Camos V. Time constraints and resource sharing in adults' working memory spans. Journal of Experimental Psychology: General 2004;133:83–100. [PubMed: 14979753]

Bentler, PM. EQS Structural Equations Program Manual. Encino, CA: Multivariate Software, Inc; 1995.

Camara WJ, Nathan JS, Puente AE. Psychological test usage: Implications in professional psychology. Professional Psychology: Research and Practice 2000;31(2):141–154.

Collette F, Van der Linden M, Salmon E. Phonological loop and central executive functioning in Alzheimer's disease. Cortex 1999;35(1):57–72. [PubMed: 10213534]

Conway ARA, Engle RW. Working memory and retrieval: A resource-dependent inhibition model. Journal of Experimental Psychology: General 1994;123:354–373. [PubMed: 7996121]

Conway ARA, Engle RW. Individual differences in working memory capacity: More evidence for a general capacity theory. Memory 1996;4:577–590. [PubMed: 8934455]

Conway ARA, Kane MJ, Bunting MF, Hambrick DZ, Wilhelm O, Engle RW. Working memory span tasks: a methodological review and user's guide. Psychonomic Bulletin and Review 2005;12:769–786. [PubMed: 16523997]

Cowan N. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behavior and Brain Sciences 2001;24:87.

Cowan, N. Working memory capacity. New York: Psychology Press; 2005.

Daneman M, Carpenter PA. Individual differences in working memory and reading. Journal of Verbal Learning & Verbal Learning 1980;19:450–466.

Daneman M, Merickle PM. Working memory and language comprehension: A meta-analysis. Psychonomic Bulletin & Review 1996;3:422–433.

Dobbs AR, Rule BG. Adult age differences in working memory. Psychology and Aging 1989;4:500–503. [PubMed: 2619956]

Engle, RW.; Kane, MJ.; Tuholski, SW. Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In: Miyake, A.; Shah, P., editors. Models of working memory: Mechanisms of active maintenance and executive control. London: Cambridge Press; 1999.

Engle RW, Tuholski SW, Laughlin JE, Conway ARA. Working memory, short term memory and general fluid intelligence: A latent variable approach. Journal of Experimental Psychology: General 1999;128:309–331. [PubMed: 10513398]

Engle, RW.; Oransky, N. The evolution from short-term to working memory: Multi-store to dynamic models of temporary storage. In: Sternberg, R., editor. The nature of human cognition. Cambridge, MA: MIT Press; 1999.

Ericsson K, Kintsch W. Long-term working memory. Psychological Review 1995;102:211–245. [PubMed: 7740089]

Fleming D, Goldberg TE, Goldman JM, Weinberger DR. Verbal working memory dysfunction in schizophrenia: Use of a Brown-Peterson paradigm. Psychiatry Research 1995;56:155–161. [PubMed: 7667440]

Gläscher J, Tranel D, Paul LK, Rudrauf D, Rorden C, Hornaday A, Grabowski T, Damasio H, Adolphs R. Lesion mapping of the cognitive abilities linked to intelligence. Neuron 2009;61:681–691. [PubMed: 19285465]

Goldman-Rakic PS. Working memory dysfunction in schizophrenia. Journal of Neuropsychiatry 1994;6:348–357.

Jamshidian M, Bentler PM. ML estimation of mean and covariance structures with missing data using complete data routines. Journal of Educational and Behavioral Statistics 1999;24:21–41.

Kail R, Hall LK. Sources of developmental change in children's word-problem performance. Journal of Educational Psychology 1999;9:660–668.

Kane MJ, Engle RW. The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual differences perspective. Psychonomic Bulletin & Review 2002;9:637–671. [PubMed: 12613671]

Kaplan, E.; Fein, D.; Kramer, J.; Delis, DC.; Morris, R. WISC-III as a process instrument. San Antonio, TX: The Psychological Corportation; 1999.

Kaufman AS, Reynolds CR, McLean JE. Age and WAIS-R intelligence in a national sample of adults in to 20- to 74- year age range: A cross-sectional analysis with education level controlled. Intelligence 1989;13:235–253.

Klein K, Fiss WH. The reliability and stability of the Turner and Engle working memory task. Behavior Research Methods, Instruments, and Computers 1999;31:429–432.

Kyllonen PC, Christal RE. Reasoning ability is (little more than) working memory capacity. Intelligence 1990;14:389–433.

Li KZH. Selection from working memory: On the relationship between processing and storage components. Aging, Neuropsychology, and Cognition 1999;6:99–116.

Oberauer K. Access to information in working memory: Exploring the focus of attention. Journal of Experimental Psychology: Learning, Memory, and Cognition 2002;28:411–421.

Pasini A, Paloscia C, Alessandrelli R, Porfirio MC, Curatolo P. Attention and executive function profile in drug naïve ADHD subtypes. Brain & Development 2007;29:400–408. [PubMed: 17207595]

Psychological Corporation. Weschler Adult Intelligence Scale-Third Edition and Wechsler Memory Scale-Third Edition technical manual. San Antonio, TX: The Psychological Corporation; 1997.

Rabin LA, Barr WB, Burton LA. Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. Archives of Clinical Neuropsychology 2005;20:33–65. [PubMed: 15620813]

Reynolds C. Forward and backward memory span should not be combined for clinical analysis. Archives of Clinical Neuropsychology 1997;12:29–40. [PubMed: 14588432]

Robert M, Savoie N. Are there gender differences in verbal and visuospatial working-memory resources? The European Journal of Cognitive Psychology 2006;18(3):378–397.

Shelton JT, Elliott EM, Hill BD, Calamia MR, Gouvier WD. A comparison of laboratory and clinical working memory tests and their prediction of fluid intelligence. Intelligence 2009;37(3):283–293. [PubMed: 20161647]

Shelton JA, Metzger RL, Elliott EM. Modified lag task as a measure of working memory. Behavior Research Methods 2007;39(3):482–493. [PubMed: 17958159]

SPSS Inc. SPSS Base 15.0 for Windows User's Guide. Chicago: SPSS, Inc; 2006.

Tabachnick, BG.; Fidell, LS. Using Multivariate Statistics. 4. Boston: Allyn & Bacon; 2001.

Tulsky, DA.; Saklofske, DH.; Zhu, J. Revising a standard: An evaluation of the origin and development of the WAIS-III. In: Tulsky, D.; Chelune, G.; Ivnik, R.; Prifitera, A.; Saklofske, D.; Heaton, R.; Bornstein, R.; Ledbetter, M., editors. Clinical Interpretation of the WAIS-III and WMS-III. San Diego: Academic Press; 2003. p. 43-92.

Turner ML, Engle RW. Is working memory capacity task dependent? Journal of Memory and Language 1989;28:127–154.

Ullman, J. Structural equation modeling. In: Tabachnick, B.; Fidell, L., editors. Using Multivariate Statistics. 4. Boston: Allyn & Bacon; 2001. p. 653-771.

Unsworth N, Engle RW. Working memory capacity and fluid abilities: Examing the correlation between Operation Span and Raven. Intelligence 2005;33:67–81.

Unsworth N, Engle RW. Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. Journal of Memory and Language 2006;54:68–80.

Unsworth N, Engle RW. The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. Psychological Review 2007;114(1):104–132. [PubMed: 17227183]

Unsworth N, Heitz RP, Schrock JC, Engle RW. An automated version of the operation span task. Behavior Research Methods 2005;37:498–505. [PubMed: 16405146]

Wechsler, D. WAIS-III administration and scoring manual. San Antonio, TX: The Psychological Corporation; 1997a.

Wechsler, D. WMS-III administration and scoring manual. San Antonio, TX: The Psychological Corporation; 1997b.

Wechsler, D. Manual for the Wechsler Adult Intelligence Scale. 4. San Antonio, TX: Pearson; 2008.

Wilde NJ, Strauss E, Tulsky DS. Memory span on the Wechsler scales. Journal of Clinical and Experimental Neuropsychology 2004;26(4):539–549. [PubMed: 15512941]

Table 1

Descriptive Statistics of WAIS-III Variables

WAIS-III Index Score or Subtest Raw Score	Mean (SD)
Full Scale IQ	110.41 (11.11)
Working Memory Index	107.16 (12.51)
Vocabulary	47.79 (7.56)
Similarities	24.58 (3.96)
Arithmetic	13.92 (3.12)
Digit Span	19.49 (3.75)
Information	18.82 (4.31)
Comprehension	22.12 (4.29)
Letter-Number Sequencing	12.53 (2.65)
Picture Completion	20.64 (3.03)
Digit Symbol Coding	88.17 (13.31)
Block Design	45.85 (10.89)
Matrix Reasoning	19.93 (2.88)
Picture Arrangement	15.57 (3.37)
Symbol Search	41.04 (6.87)

Note. Only Full Scale IQ and Working Memory Index are standardized scores. All others are raw scores.

Table 2

Descriptive Statistics of Experimental Cognitive Variables

Cognitive Tasks	**Mean (SD)**
Listening Span Score	29.70 (12.28)
Operation Span Score	44.13 (15.78)
Operation Span Errors	5.58 (2.94)
Modified Lag Task Score	54.41 (16.50)

Note. All scores are raw scores.

Table 3

Working Memory Criterion Construct Factor Loadings

Experiment Cognitive Task	Factor Loading	Initial Communality
Listening span	.81	.39
Operation span	.70	.35
Modified lag task	.59	.26
Eigenvalue	1.49	
Percent of variance explained	49.75	

Table 4

Backward Elimination Multiple Regression Analysis Final Model for WAIS-III Subtests Predicting WMCC Score

WAIS-III Subtest	***B***	***SEB***	**β**	***p* value**
Constant	−4.39	.45	-	< .001
Digit Span	.08	.02	.35	< .001
Letter-Number Sequencing	.08	.02	.23	.002
Matrix Reasoning	.05	.02	.17	.004
Vocabulary	.02	.01	.14	.024

Note. $R^2 = .43$; $F(4,175) = 33.26$, $p < .001$

Table 5

Percent of Variance Accounted for in the WMCC by Each WAIS-III Subtest Individually

WAIS-III Subtests	R^2	***p* value**
Vocabulary	.15	<.001
Similarities	.03	.031
Arithmetic	.14	<.001
Digit Span	.33	<.001
Information	.12	<.001
Comprehension	.03	.013
Letter-Number Sequencing	.28	<.001
Picture Completion	.01	.324
Digit Symbol Coding	.01	.232
Block Design	.14	<.001
Matrix Reasoning	.10	<.001
Picture Arrangement	.08	<.001
Symbol Search	.02	.081
Full Scale IQ	.27	<.001