

Examining the Incremental Validity of Working Memory
for Predicting Learning and Task Performance: A Partial Mediation Model

Devin Matthew Carter

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in
partial fulfillment of the requirements for the degree of

Master of Science
In
Psychology

E. Scott Geller, Co-Chair
Neil M. A. Hauenstein, Co-Chair
Roseanne J. Foti

October 6th, 2017
Blacksburg, Virginia

Keywords: Working Memory, Intelligence, Knowledge, Learning, Performance

Examining the Incremental Validity of Working Memory for Predicting Learning and Task Performance: A Partial Mediation Model

Devin M. Carter

ACADEMIC ABSTRACT

General intelligence (“g”) has long been used as an effective predictor of both learning and job performance. Further, other more specific cognitive abilities have not been able to consistently predict incremental variance in job knowledge and job performance beyond “g”. However, the processes associated with working memory (WM) are important for these outcomes and are not captured by our traditional tests of “g”. This study tested a partial mediation model in which working memory (WM) incrementally predicts task performance above “g” through task knowledge and through a direct effect. Participants were given measures of “g” and WM in a lab. They were then given a learning opportunity and a task that applies this newly learned knowledge in order to test the effects of WM. Results indicate that WM explains additional variance in both task knowledge and task performance, and the partial mediation model was supported using one of the two WM tasks used.

Examining the Incremental Validity of Working Memory for Predicting Learning and Task Performance: A Partial Mediation Model

Devin M. Carter

GENERAL AUDIENCE ABSTRACT

General intelligence is widely used in personnel selection because it is consistent in predicting the job performance of future employees. Other cognitive abilities have also been examined to determine whether they are able to predict job performance as well as general intelligence. However, most of these other cognitive abilities have come up short. This study hypothesized that working memory (WM) is a cognitive ability that may be able to predict job performance even after controlling for general intelligence. A sample of undergraduates completed tasks that measured general intelligence and WM, and this study examined how well each measure predicted both learning and performance on a relatively novel task. Results indicated that WM was able to predict both learning and performance after controlling for general intelligence.

Contents

1 Introduction	1
1.2 Working Memory	3
1.3 WM and g	5
1.4 Criterion-related validity for WM: Learning	6
1.5 Criterion-related validity for WM: Task Performance	7
1.6 Capacity vs. Processing Efficiency of WM	8
1.7 Overview	12
2 Literature Review	13
2.2 Measuring Intelligence	15
2.3 Practical Consequences of Using General Cognitive Ability Tests	17
2.4 Structure of WM	17
2.5 WM and Learning	24
2.6 WM and Task Performance	26
2.7 Working Memory: Capacity and Processing	27
2.8 The Current Study	30
3 Methods	33
3.1 Participants	33
3.2 Procedure	35
3.3 Measures	39
3.4 Analysis	40
4 Results	41
4.1 Descriptives and Correlations	41
4.2 Residualizing Outcome Variables for Mediation Tests	43
4.3 Baseline Mediation Model Not Controlling for Gf	44
4.4 Mediation Model Controlling for Gf	45
4.5 Exploring Both WM and Gf as Antecedents	48
4.6 Summary	49
5 Discussion	50
5.1 Conceptual Implications	50
5.2 Domain Generality vs. Specificity of WM	53
5.3 Practical Implications	54
5.4 Limitations and Future Directions	55
5.5 Conclusions	57

List of Figures

Figure 1. The direct and indirect effects of WM on task performance.	9
Figure 2. Carroll's (1993) three-stratum theory.	15
Figure 3. The most recently proposed version of Baddeley's (2000) working memory system.	18
Figure 4. An excerpt of Schnedier & McGrew's (2012) CHC model v2.1.	21
Figure 5. Beta coefficients for the model using the RSPAN & OSPAN, not controlling for Gf	44
Figure 6. Beta coefficients for the model using the RSPAN & OSPAN, controlling for Gf.	45

List of Tables

Table 1. Descriptive Statistics	41
Table 2. Correlation matrix of measured variables.	42

Chapter 1

Introduction

General cognitive ability tests have been used for personnel selection in the United States since World War I (Driskell & Olmstead, 1989; Thomas & Scroggins, 2006). General cognitive ability is the strongest predictor of training performance (Ghiselli & Brown, 1951; Gottfredson, 1997; Hunter, 1986) as well as job performance across job families (Ghiselli, 1973; Hunter, 1986; Schmidt, 2002) Furthermore, it is widely accepted that job knowledge mediates the relationship between cognitive ability and task performance (Hunter, 1986; Palumbo, Miller, Shalin, & Steele-Johnson, 2005; Schmidt, Hunter, & Outerbridge, 1986); individuals of higher intelligence are more likely to obtain higher levels of job knowledge which, in turn, leads to better job performance. Schmidt and Hunter (2004) found that the relationship between cognitive ability and job performance is moderated by the complexity of the job, such that cognitive ability has higher predictive validity for job performance for more complex jobs and lower predictive validity for less complex jobs. The explanation given by Schmidt and Hunter is that complex

jobs require more complex and integrated knowledge than simple jobs, and more intelligent people acquire complex knowledge efficiently.

Cognitive ability is hierarchically structured in most intelligence theories, such that general intelligence (“g”) is reflected by specific abilities (“s”) (Spearman, 1904). Researchers have examined whether measurement of specific cognitive abilities “s” provides incremental validity over “g” when predicting job performance. Common specific abilities examined include mathematical knowledge, reading comprehension, and spatial skills (Carroll, 1993). Although there is evidence that perceptual speed and spatial abilities predict performance beyond “g” for certain occupations (Mount, In-sue, & Burns, 2008), the preponderance of evidence suggests that s-abilities fail to provide incremental validity over “g”. Indeed, Ree and Earles (1991) found that across 82 jobs and over 78,000 participants in the U.S. Air Force, specific abilities captured by the Armed Services Vocational Aptitude Battery (ASVAB) did not explain unique variance in the prediction of training success. This finding was replicated in that specific abilities measured by the ASVAB were not found to predict above and beyond “g” on work sample tests (Ree, Earles, & Teachout, 1994). Other researchers have also found that s-abilities explain little incremental variance to the prediction of both training criteria (Ree & Earles, 1991; Stanhope & Surface, 2014) and job performance (Ree, Earles, & Teachout, 1994; Schmidt & Hunter, 2004).

However, measures of working memory (WM) have not been included in the mix of s-abilities that are usually studied. Although WM is sometimes overlooked because it is an s-ability, cognitive processes reflected by WM are essential for both learning and task performance. These outcomes are both important for personnel selection researchers and practitioners. In the limited studies conducted using WM, WM scores have been shown to predict incremental variance above “g” for both learning outcomes and task performance, as

discussed later. Thus, the first purpose of this study is to add to the limited literature and provide additional evidence that WM scores incrementally predict learning outcomes in addition to task performance.

It is noteworthy that the current study cites literature on task knowledge, job knowledge, task performance, and job performance. Although task knowledge can be slightly distinct from job knowledge, the current study posits that it is intuitive to infer that the same processes occur to attain both task knowledge and job knowledge. Therefore, when considering underlying theory, task knowledge and job knowledge are treated as the same. The same can be said for task performance and job performance.

Past research has examined WM as an incremental predictor using both knowledge and performance as criterion measures; however, past studies have failed to incorporate both. Thus, this study attempts to merge past findings which support WM's criterion-related validity into a more complete model.

Working Memory

A model of the working memory system was first proposed by Baddeley and Hitch (1974) and models of working memory remain popular in cognitive psychology (Baddeley, 2000; Miyake & Shaw, 1999; Repovs & Baddeley, 2006). Baddeley and Hitch's (1974) definition, as well as much of the associated literature, propose WM consists of many processes including: 1.) *maintaining* the activation of goal-related stimuli, 2.) *updating* representations of stimuli with the presentation of new or changing stimuli, 3.) *inhibiting* interference from distracting stimuli, 4.) *task-switching* in order to manipulate new or irrelevant information while retaining previously attended stimuli, and 5.) *rehearsing* goal-relevant information in order to avoid forgetting (Kane, Conway, Hambrick, & Engle, 2007).

Measurement of WM. Because of the complexity and multiple cognitive processes included in WM, researchers do not claim to have measures that fully map the WM construct. Kane et al. (2007) pointed out that different measures of WM may require the use of different WM processes to some degree. For example, the Automated Operation Span (OSPAN) task presents participants with simple math problems. Between each of these math problems, participants are presented with a single letter to be remembered on the screen. After three to seven math problems and letters, participants report the letters, in the order presented, on the subsequent screen. Their success is based on their ability to *maintain* the letters in memory, *update* the sequence of letters as each one is added, *inhibit* the interference of the math problems in remembering the letters, and *task-switch* in order to manipulate the newly presented quantitative information, all while *rehearsing* the sequence of goal-relevant letters in order to recall them on the subsequent screen. Although a simple task, the OSPAN taps into all of these processes (Kane et al., 2007) and is therefore one of the most widely used WM tasks in the research literature.

The *N*-back task is also popular, whereby participants are presented abstract shapes on a computer screen, one at a time and with a short exposure rate. Participants must click the mouse every time a shape is presented that is identical to the shape presented *N* shapes ago. For example, if the task is set at a three-back rule, the participant must click the mouse if the shape that is shown is identical to the shape shown three shapes back. Success on this task clearly requires maintaining, updating, inhibiting, and rehearsing information. However, it involves less task-switching than the OSPAN because there is no separate task competing for the participants' attention. Thus, different measures of WM can require the utilization of different processes.

Multiple studies have supported the predictive accuracy of WM tests for learning criteria. For example, WM was positively associated with the ability to learn a second language (DeKeyser & Koeth, 2011; Hummel, 2009; Linck & Weiss, 2015; Wen & Skehan, 2011). Also, Gersten, Jordan, and Flojo (2005) found that WM scores predicted mathematical ability in children's first year of formal schooling, a time when the learning curve is steep. WM scores have been shown to be positively correlated to learning computational skills (Wilson & Swanson, 2001), learning mathematical word problem solving skills (Swanson & Sachse-Lee, 2001), learning visually-presented material (Logie, Gilhooly, & Wynn, 1994), and remembering lengthy instructions (Gathercole, Lamont, & Alloway, 2006). Lastly, evidence shows us that individuals' WM scores are positively associated with a better memory of something they have just been shown, keeping track of where they are in the process of completing a learned task, and remembering more specific details of what they have learned and how learned information can be applied in different settings (Alloway, 2009).

WM and g

The development of WM measures and the examination of the cognitive processes involved in WM have led researchers to integrate WM into existing models of "g". McGrew (2005) combined Carroll's (1993) and Cattell's (1971) models of "g" to create the most widely accepted model of intelligence today – the Cattell-Horn-Carroll (CHC) model (Kaufman, 2009). In this model, psychometric "g" is a third-order construct (Stratum III) representing overarching general cognitive ability.

The second-order constructs in Stratum II consist of broad abilities such as fluid reasoning, quantitative knowledge, reading and writing, comprehension knowledge, short-term memory, long-term storage and retrieval, visual and auditory processing, and processing speed.

Because of WM's theoretical links to short-term memory (discussed later), WM has been treated as a first-order Stratum I ability hierarchically lower than short-term memory that reflects short-term memory capabilities (Schneider & McGrew, 2012). The proposed study adopts the CHC model of intelligence and considers WM to be a specific, first-order ability in the structure of general intelligence.

Although WM is commonly considered to be part of "g", most widely-used general cognitive ability theories were published prior to the development of the WM arguments. Most cognitive ability tests measure either fluid reasoning (e.g., Raven's Progressive Matrices) or comprehension knowledge (e.g., the ASVAB) along with some other specific abilities that typically do not include accepted WM measures (Conway, McNamara, & Engel, 2013). It should be noted that when I refer to studies that measure "g", I am referring to studies that use g-measures that do not include measures of WM.

These frequently-used assessments of "g" have been shown to predict various learning outcomes (Filičková, Ropovik, Bobaková, & Kovalčíková, 2015; Ren, Schweizer, Wang, & Xu, 2015; Swanson & Kim, 2007) as well as academic achievement (Kuncel & Hezlett, 2007; Sackett, Borneman, & Connelly, 2008). Given these common findings, it is worthwhile to examine the literature with regard to how well WM predicts similar learning outcomes above and beyond measures of g.

Criterion-related validity for WM: Learning

The previously mentioned studies that used WM scores to predict learning outcomes (e.g., learning a second language, learning computational skills, learning visually presented material, and remembering instructions) did not control for "g". However, Alloway (2009) found that among children with learning disabilities, WM scores predicted learning criteria better than

“g”. Furthermore, Alloway and Alloway (2010) found that WM measures predicted learning better than “g” scores over the course of six years of schooling in young children. This is consistent with Luo, Thomson, and Detterman’s (2006) findings that WM explained additional variance in school-aged children’s scholastic achievement above a composite of fluid reasoning. Finally, Krumm, Ziegler, and Buehner (2008) found that a WM measure of verbal storage and processing added incremental validity to the prediction of language learning in undergraduate college students.

It should be noted that much of the evidence supporting differences in WM as an antecedent of learning examines learning in the classroom with school-aged children. The present study examined these differences in WM as a predictor of adult learning.

Criterion-related validity for WM: Task Performance

Given that research has consistently shown g-scores to be the most valid predictors of task performance (Hunter & Hunter, 1984; Ree & Earles, 1992; Schmidt, 2002), and that other specific abilities add little explanatory power (Jensen, 1986; Olea & Ree, 1994; Ree et al., 1994), there is limited research on the possible incremental validity provided by WM scores. However, three studies have addressed this issue.

Wolfe, Alderton, Larson, and Held (1995) found when adding a battery of WM, spatial abilities, reasoning and perceptual speed to the ASVAB, the prediction of task performance in Navy Technical schools increased by up to 16.7%. Independent from the rest of the battery, WM predicted a significant amount of variance in performance in one-third of the schools in the study, including up to 5% additional variance predicted in task performance above the ASVAB in one of the schools examined. Although a small percent of predicted variance in absolute terms, Schmidt, Hunter, and Dunn (1987) calculated that as low as a 3% increase in the

predictive validity of the ASVAB could save the Navy \$83 million every year (Wolfe et al., 1995). Additionally, Held and Wolfe (1997) found that a WM test added 5%, 10%, and 5% validity to the prediction of performance among candidates on basic Air Force training modules, advanced Air Force training modules, and Operations Specialist ratings, respectively.

Bosco, Allen, and Singh (2015) used two measures of WM to predict task performance above and beyond a commonly used measure of general cognitive ability – the Wonderlic Personnel Test (WPT). Two of their four studies measured undergraduates' WM and presented them with a supply chain management simulation task. Their other two studies used WM scores to predict supervisory ratings of bank employees. Across four studies conducted with both undergraduates and employees in the workforce, they found that WM scores predicted task simulation performance and supervisory ratings significantly better than the WPT. They called for future research to aim for a better understanding of the relationship between their measures of WM and task/job performance.

These studies are among the limited studies conducted using WM to predict task performance. The reason for the minimal studies evaluating WM as an incremental predictor is due to the fact that many specific abilities have not been supported as incremental predictors in past literature. However, I propose the cognitive processes and efficiency associated with WM are important for both the learning of task knowledge and the successful application of this knowledge. Lastly, it's noteworthy that almost none of the cognitive ability assessments capture these specific WM cognitive processes. The cognitive processes outlined above result in two central functions of WM.

Capacity vs. Processing Efficiency of WM

The second purpose of this study is based on the argument that WM represents both structural and processing functions (Baddeley and Hitch, 1974), and these functions likely affect task knowledge and task performance differentially. The structural argument refers to the capacity of WM (WMC) to maintain information in the presence of distracting information; the WM processing (WMP) argument refers to a person's ability to update, manipulate, and retrieve learned information, while inhibiting competing, distracting goals. I argue that the structure/processing distinction results in a more complex causal model than learning fully mediating WM and task performance. Rather, learning will only partially mediate the WM-task performance relationship because the WMP efficiency function will also directly affect task performance. The second purpose of this study, therefore, is to empirically test this partially mediated model (see Figure 1).

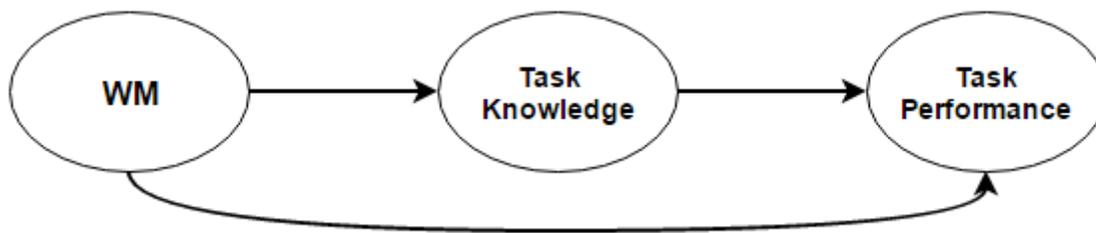


Figure 1. The direct and indirect effects of WM on task performance.

WM – A mediated model argument. In order to add to the limited research on working memory as a contributor to the prediction of learning task knowledge and task performance above “g”, the proposed study aims to examine WM, learning, and task performance from this dual functionality perspective. As described earlier, the traditional mediated model poses that general cognitive ability positively affects the acquisition of job knowledge during training, and the acquisition of job knowledge positively affects task performance on the job (Hunter, 1986).

Given the studies that have provided evidence that WM is an important predictor when examining learning outcomes, I suggest WM will incrementally predict learning outcomes which, in turn, incrementally predict task performance through the mediation of learning. The proposed study includes both learning and task performance as outcomes in order to paint a more complete picture of partial mediation.

When WM is viewed only in terms of capacity, task knowledge fully mediates WM and task performance. As stated earlier, WMC is the *amount* of information a person can hold in memory while being exposed to irrelevant information (Daneman & Carpenter, 1980). One could posit that without considering WMP, WMC could be a part of a fully mediated model, such that task knowledge fully mediates the relationship between WM and task performance because the storage processes associated with WMC are important for learning to take place. I extend this perspective and adopt the dual functionality approach, where both WMC and WMP contribute to the learning of information. Thus, the indirect effect of WM on task performance through the learning of task knowledge will be tested. However, this study adds another viewpoint, incorporating the processing efficiency perspective of WM and proposing a partially mediated model (Figure 1).

WMP – A processing efficiency argument. It is generally accepted in the literature that more intelligent people can gather and retain knowledge more quickly than less intelligent people. As task experience is gained, though, less intelligent people likely attain similar levels of job knowledge as more intelligent people, especially for simple jobs, where less job knowledge is required and the tasks are relatively simple (cf. Schmidt & Hunter, 2004).

Although a less intelligent person is likely to eventually reach similar levels of task knowledge as a more intelligent person, I propose a person with a better WM will still perform

better on a task than a person with worse WM. That is, people with better WM have better WMP and will achieve higher levels of performance given their advantages in processing efficiency (i.e., their ability to accurately and quickly store information, manipulate new information, and block out irrelevant distractors). Therefore, even though both better WM people and worse WM people may possess comparable knowledge to perform a given task, people with better WM will be better at attending to specific cognitive processes, which results in better performance.

This is consistent with the dual perspective view of WM described by Baddeley and Hitch (1974). The capacity facet of the WM system is offset by knowledge of the task that has been learned, but WMP allows one person to more efficiently recall information, manipulate information while holding relevant information in memory, and inhibit distracting or irrelevant stimuli while performing a relevant task. This framework is important because it supports a partially mediated model (Figure 1) where WM predicts task performance directly, independent of learned task knowledge. The processing efficiency perspective of WM supports the notion that when people are performing a task, people with better WM will perform better due to their ability to carry out the important cognitive processes involved in successfully completing the task.

Bosco et al. (2015) provide some evidence for WM as a predictor of task performance in that their measures of WM predicted both simulation performance and supervisor ratings of on-the-job performance. In line with their findings, the proposed study suggests that WM directly affects task performance because WM reflects cognitive processes that are necessary for efficiently completing a task; therefore, WM measures tap into processes that may uniquely predict task performance. Therefore, the present study posits that WM will have both an indirect and direct effect on task performance when controlling for “g”.

Overview

Past studies have examined the predictive accuracy of WM scores for both learning criteria and task performance. In this study, two different WM tasks – the Automated Operation Span (OSPAN) and the Automated Reading Span (RSPAN) – was used to examine the indirect effect of WM on task performance above “g”, as mediated by task knowledge. Furthermore, this study examined the direct effect of WM on task performance beyond “g”.

Participants (undergraduate students) in the current study were given commonly-used measures of Gf and WM, and tasked with learning material on the specifics properly-formatting an APA style reference page via an online tutorial. An APA style knowledge pretest and posttest assessed knowledge gains from the tutorial. Lastly, students were tasked with compiling a properly-formatted APA style reference page based on knowledge they have learned from the tutorial.

The two primary goals for the current study were: 1.) to provide additional evidence of the incremental validity of WM scores over “g” for both learning and task performance, and 2.) to empirically test the model which suggests task knowledge partially mediates the effects of WM on task performance (Figure 1).

Chapter 2

Literature Review

Research on cognitive ability has been conducted since astronomer Bressel found individual differences in choice reaction time in the early 1800s (Carroll, 1993). Numerous theories of cognitive ability have emerged over the past 100 years.

In the organizational sciences “intelligence,” “general mental ability,” “cognitive ability,” “psychometric g,” “Gf,” and “g” are treated as interchangeable. The classification of cognitive ability as one overarching construct stems from Charles Spearman’s (1904) early research with cognitive tests. Spearman assessed school-aged children’s performance on several mental tests including numerical ability, verbal fluency, and spatial manipulation (Kaufman, 2009). He found that children who scored relatively high on one cognitive test tended to score high on other cognitive tests. Spearman coined the term “general factor” or “g” to explain this positive manifold among tests, and this positive manifold is one of the most replicated findings in the field of psychology (Deary, 2000). Psychometric “g”, therefore, refers to the performance of

individuals across a representative set of cognitive tests (Jensen, 2002). Spearman also recognized a second lower level factor, “s”, which influences the ability to perform on specific cognitive tasks.

Building on Spearman’s model, Horn and Cattell (1966) identified two facets of intelligence, coining the terms fluid intelligence, or Gf, and crystallized intelligence, or Gc. Gf refers to the ability to reason, think abstractly, and using higher order processes to solve novel problems (Cattell, 1987), while Gc refers to the knowledge a person has gained through culture, education, experiences, and knowledge gained through “investment” of fluid intelligence resources (McGrew, 2005). Cattell and Horn later extended Gf-Gc theory to include other broad abilities such as fluency, short-term apprehension and retrieval, and correct decision speed, among others (McGrew, 2009) which are incorporated into the most widely referenced model of human intelligence—CHC theory.

John Carroll (1993) sought to develop an inclusive framework of cognitive ability by reanalyzing over 400 studies of cognitive ability. Carroll’s (1993) resulting book from this study has been referred to by some as a “periodic table” for intelligence research (Horn, 1998, p. 58). Carroll concluded that cognitive ability had three levels, or strata. Stratum III is overarching general intelligence, or what Carroll argued to be Spearman’s psychometric “g”. Stratum II are “broad abilities” including fluid intelligence (Gf), crystallized intelligence (Gc), general memory and learning (Gy), visual perception (Gv), auditory perception (Gu), retrieval ability (Gr), cognitive speediness (Gs), and processing speed (Gt).

Below broad abilities are over 80 stratum I specific abilities; for example, narrow abilities for fluid reasoning include general sequential reasoning, induction, quantitative reasoning, and

Piagetian reasoning. The general structure of Carroll's three-stratum theory is presented in Figure 2.

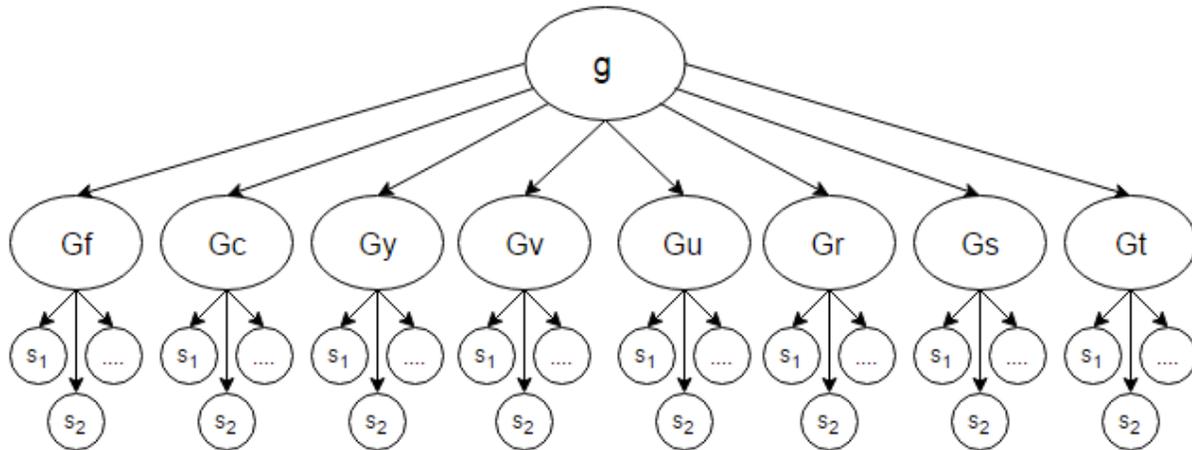


Figure 2. Carroll's (1993) three-stratum theory. Broad abilities are fluid intelligence (Gf), crystallized intelligence (Gc), general memory and learning (Gy), broad visual perception (Gv), broad auditory perception (Gu), broad retrieval ability (Gr), broad cognitive speediness (Gs), and processing speed (Gt). There are over 80 specific abilities (e.g., s_1 , s_2 , etc.) hierarchically below the broad abilities.

Carroll recognized the similarity between his three-stratum theory and Horn's (1991) model of general intelligence. Horn did not argue for a general higher-order intelligence that influences abilities at the lower strata, whereas Carroll argued that an overarching "g" construct must be included in any theory of intelligence because of the overwhelming empirical evidence (Flanagan & Dixon, 2013). Although there is still some debate regarding a single "g" factor, the CHC model includes psychometric "g" as the highest order factor of cognitive ability.

Measuring Intelligence

Cognitive ability or psychometric "g" has been highly regarded as the most accurate predictor of job performance (Judge, Klinger, & Simon, 2010; Kuncel, Hezlett, & Ones, 2004; Salgado, Anderson, Moscoso, Bertua, de Fruyt, & Rolland, 2003; Schmidt & Hunter, 1998).

There are numerous tests of cognitive ability; some of the most widely cited cognitive measures

used in the organizational sciences include the Wonderlic Personnel Test (WPT), Raven's Advanced Progressive Matrices (RAPM; Raven, Raven, & Court, 1998), and the Armed Services Vocational Aptitude Battery (ASVAB) (Schmidt & Hunter, 2004; Talboy, 2011).

The WPT is a commercial test based on the concept that scores from shorter cognitive tests approach the levels of reliability and validity found in scores from longer assessments of cognitive ability (Foster, Shipstead, Harrison, Hicks, Redick & Engle, 2015), as it typically takes about 12 minutes to complete. The WPT contains questions involving analogies, spatial reasoning, arithmetic, similarity problems, and word definitions. Therefore, it addresses the broad abilities of quantitative reasoning (Gq), comprehension knowledge (Gc), and fluid reasoning (Gf) (Matthews & Lassiter, 2007).

Participants taking the RAPM examine a matrix of shapes, lines, and patterns, and determine which figure is missing from one of the fields in the matrix. Test takers must use abstract reasoning and logic to find a specific rule in a given matrix and use this rule to fill in the blank space. The RAPM has been shown to be the most valid indicator of fluid reasoning (Gf) ability (Chuderski, 2014; Gignac, 2014). The ASVAB is given to military applicants and contains subtests that address paragraph and word comprehension, and arithmetic reasoning and knowledge (ASVAB, 2016). Thus, Gf, Gc, and Grw broad abilities are captured by the ASVAB.

It is clear that Gf is at the center of many cognitive ability tests. Over the past few decades, researchers have found Gf to be the highest correlate to psychometric "g" (Carroll, 1993, 2003; Conway, Jarrold, Kane, Miyake & Towse, 2007; Undheim, 1981; Undheim & Gustafsson, 1987). This is not surprising given that intelligence is often defined as Gf (Carroll, 1993; Spearman, 1927), and reported Gf-g correlations are sometimes so strong that researchers

conclude *G_f* is isomorphic with general intelligence (Chuderski, 2013; Schneider & McGrew, 2012).

Practical Consequences of Using General Cognitive Ability Tests

The observed predictive accuracy of cognitive ability scores range from 0.2 to 0.35, and corrected meta-analytic estimates place the values closer to 0.4 to 0.5 (Hunter & Hunter, 1984; Pearlman, Schmidt & Hunter 1980; Schmidt & Hunter, 2004; Viswesvaran & Ones, 2002). The predictive validity of cognitive ability scores routinely surpasses other predictor of job performance, including interviews, personality tests, and biodata (Ghiselli, 1973; Hunter, 1986; Ones et al., 2012; Schmidt, 2002). Cognitive ability is also linked to other job-related outcomes, such as success in training (Barret, Polomsky, & McDaniel, 1999; Hirsh, Northrup, & Schmidt, 1986), job satisfaction (Ganzach, 1998), leadership (Judge, Colbert, & Ilies, 2004; Lord, De Vader & Alliger, 1986), creativity (Kuncel, Hezlett, & Ones, 2004) and counterproductive work behaviors (Judge, Klinger, & Simon, 2010).

Because of the predictive validity of *g*-scores for so many outcomes and the relatively little support for specific abilities in the incremental prediction above “*g*”, researchers tend to ignore specific abilities. They assume “*g*” to captures the variance associated with the abilities lower in the hierarchy. However, the present study posits that the processes associated with WM are important, and are not captured by measures of “*g*”.

Structure of WM

The construct of WM was first used by Miller, Galanter, and Pribram (1960) when referring to Miller’s (1956) idea of chunking information for memory purposes. However, the role of WM in relation to other cognitive processes wasn’t thoroughly explored until Baddeley and Hitch (1974) proposed a WM system model. Baddeley’s (2000) more recent revision of the

WM system includes four components: 1) the phonological loop, 2) the visuo-spatial sketchpad, 3) the episodic buffer, and 4) the central executive (see Figure 3).

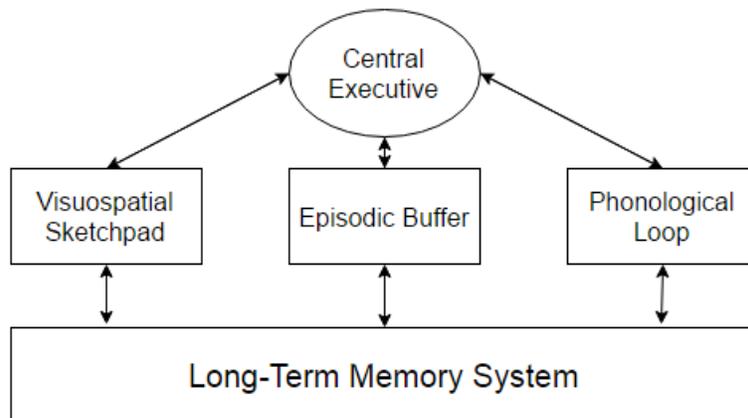


Figure 3. The most recently proposed version of Baddeley's (2000) working memory system.

The phonological loop is a hypothetical entity in the information processing system that stores auditory information and maintains this information via rehearsal (Baddeley, 2000). When people perceive an auditory stimulus, their ability to maintain this raw information is extremely limited (Darwin, Turvey, & Crowder, 1972; Guttman & Julesz, 1963; Sams, Hari, Rif, & Knuutila, 1993). However, the phonological loop has a rehearsal mechanism that allows people to store this information for a longer period of time by renewing the presence of the auditory information.

For example, imagine a person (for ease of reading, let's say this person is female) checks her voicemail while driving. She checks her voicemail, and someone needs her to call him back, and he provides her with his number. She does not have access to a pen and paper, so she has to listen to the voicemail to hear the number again, hang up the phone, and dial the caller's number from memory. More than likely, she will be repeating the numbers in her head, time after time, until she is able to dial the number. The ability to accurately dial the number is a

result of the phonological loop, and more specifically, the articulatory rehearsal component of the phonological loop, which is responsible for the rehearsal of auditory information while it is in the WM system (Baddeley, 1986). The phonological loop has gathered much supporting evidence for its existence since it was first proposed (Bloom & Watkins, 1999; Duffau, Gatignol, Denvil, Lopes, & Capelle, 2003; Wilson, 2001).

The visuo-spatial sketchpad is responsible for object and spatial short-term memory. Object short-term memory maintains what the object “is”, as well as other visual properties such as color, shape, and texture (Baddeley, 1997; Baddeley, Eysenck, & Anderson, 2009). Spatial short-term memory allows individuals to keep a snapshot of where an object is in relation to other visuo-spatial representations (i.e., the location of an object in space). These two components of the visuo-spatial sketchpad can work independently or together because of their different paths from sensation to perception in the brain (e.g., Denis, Logie, & Cornaldo, 2012).

The episodic buffer is the mechanism by which information in the visuo-spatial sketchpad and information in the phonological loop are integrated into a single event or episode (Baddeley, Allen, & Hitch, 2011). For example, when a person watches a cartoon on TV, immediately after a certain scene, (s)he could probably describe or even draw the couple seconds of the scene that just occurred, including where each character was in relation to others, the color clothes the characters were wearing, what the characters last said to each other, and how the visuals changed with the audio in the scene with accurate chronology. This is a phenomenon explained by the presence of an episodic buffer. The episodic buffer was added to the model by Baddeley (2000) in order to explain this integration of the phonological loop and visuo-spatial sketchpad as well as to serve as a link between these WM components and long-term memory.

Baddeley (2000) refers to the three WM system components as “slave systems” because these components are capable of only very short-term storage of different types of information unless they are maintained by the central executive component of the WM system. Just as a senior executive of an organization supervises his/her employees throughout the day, the central executive supervises these three components and coordinates the information of all three in order to make sense of input information from three different main sources (Henry, 2012). The central executive is also responsible for shifting among these slave systems in order to perform goal-related tasks. Furthermore, the central executive function is to maintain needed information in the slave systems while attending to new stimuli, whether this new stimuli is auditory, visuo-spatial, or stimuli from other senses.

The central executive is also responsible for inhibiting external stimuli that may intrude upon the information in the phonological loop, visuo-spatial sketchpad, and the episodic buffer (Kane & Engle, 2002). The central executive typically prioritizes tasks to focus on the processes that will be most demanding for the WM system (Baddeley, 1996).

Baddeley (2012) has since proposed a “speculative model” (p. 22) where information first enters the visuo-spatial sketchpad and phonological loop before being integrated by the episodic buffer. Although this flow of information is intuitive, no empirical research has been conducted to examine the order in which information is used by these different WM components. Therefore, the model used in the current study (see Figure 3) represents the model most commonly used in WM research.

WM in the CHC model. When Carroll (1993) first proposed his three-stratum theory, WM was not included most likely because only six of more than 400 studies reviewed included measures of WM (Ackerman et al., 2005). Carroll, however, pointed out the apparent importance

of a “central processor” for reasoning and other types of cognitive tasks (p. 646). When mentioning this central processor, he is essentially referring to Baddeley’s (1986) model of WM. Therefore, Carroll was referencing the WM system as an overarching system that affects a wide range of cognitive processes. Carroll mentioned that more measures of WM needed to be developed, and their psychometric properties needed to be established before WM can be fully understood.

The increase in the study of WM across many fields (e.g., Kieras et al., 1999; Miyake & Shah, 1999; O’Reilly et al., 1999; Vallar, 2006) and the development of new tests have resulted in WM being incorporated into the CHC model. Most current versions of the CHC model include WM as part of the broad ability of Gsm (Stratum II, See Figure 4), or short-term memory because of theoretical links between short-term and WM, as well as spatial and object short-term memory being part of the WM system.

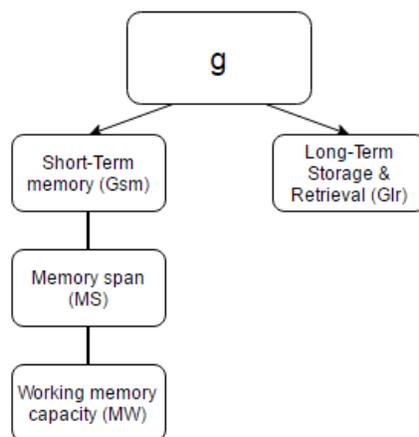


Figure 4. An excerpt of Schnedier & McGrew’s (2012) CHC model v2.1. WM is considered a specific ability of the broad ability of Short-Term Memory (Gsm).

Relationship between WM and STM. Researchers often consider WM as synonymous with short-term memory (STM; e.g., Aben, Stapert, & Blokland, 2012). Although these

constructs are strongly related both theoretically (Baddeley & Hitch, 1974; Daneman & Carpenter, 1980) and empirically (Aben, Stapert, Blokland, 2012; Engle, Tuholski, Laughlin, & Conway, 1999), there are key differences. As explained by Carroll (1993), STM is the capacity to hold small amounts of information to be recalled in a short period of time. Miller (1956) proposed the well-known 7 ± 2 bits of information that can be held in STM before individuals start to fail to recall information. Many measures of STM involve memory span tasks that require memorization of number, letters, words, shapes, or objects to be recalled. The pieces of information that can be successfully recalled reflect a person's short-term memory capacity.

The construct of WM, on the other hand, involves an attentional aspect that maintains information in the short-term stores (i.e., the phonological loop, the visuospatial sketchpad) but inhibits the influence of new, irrelevant information presented (Kane & Engle, 2002). These maintenance and attentional functions have been given many labels such as working memory capacity, executive attention, supervisory attention system, anterior attention system, controlled attention, and the shared variance among different WM tasks (Engle et al., 1999; Engle & Kane, 2004). Although STM is foundational to WM, the central executive in effectively maintaining and focusing on this information in short-term memory while blocking out interference from other task-irrelevant sources is what differentiates STM from WM. As Kane et al.(2007) stated :

“when we use the term *working memory capacity*, we refer to the attentional processes that allow for goal-directed behavior by maintaining relevant information in an active, easily accessible state outside of conscious focus, or to retrieve information from inactive memory, under conditions of interference, distraction, or conflict (p. 23)”.

WM is typically measured by a more complex version of the simple span tasks (Kane, Conway, Hambrick, & Engle, 2007) where the numbers, letters, words, or shapes to be recalled are learned while performing a competing task. For example, for the complex operation span task (OSPAN), participants are presented with a task sequence of learning a letter, solving a simple math problem, and learning another letter. The number of letters in each block can vary and the participant tries to recall all the letters presented at the end of the block. The introduction of a secondary task (i.e., the math problems) requires participants to hold the letters that were presented in short-term memory while using cognitive processes elsewhere. Even though participants try to solve these simple math problems, they are not meant to recall this information, and even must cognitively rehearse the *letters* while solving these problems in order to recall the correct letters.

The strong correlation between WM and STM is thought to occur for two reasons: 1) the WM construct consists of the short-term functions that store the information plus an executive attention or attentional control aspect, and 2) WM and STM tasks are extremely similar, differing only in the presence or absence of a secondary task that includes or excludes the executive function from the cognitive processes involved (Kane et al., 2007).

The WM – Gf relationship. Many studies in the WM literature attempt to better understand the association between WM and Gf. This is perhaps because of the mixed correlational results that are found between the two constructs. Kyllonen and Christal (1990) found a 0.80 - 0.90 correlation between performance on fluid reasoning tasks and performance on WM tasks across four studies. This relationship has been supported in some studies: Stauffer, Ree, and Caretta (1996) found a 0.995 correlation between the two, along with studies that have suggested WM and Gf to be isomorphic (Blair, 2006; Engle, 2002; Jensen, 1998).

However, the findings from numerous studies have left researchers skeptical of the isomorphic relationship between WM and Gf. Meta-analyzing correlations from 86 studies, Ackerman, Beier, and Boyle (2005) found a mean correlation of 0.48. Babcock (1994) found a similar magnitude of the relationship ($r = 0.55$) when correlating WM measures to Gf. Several others researchers have found the correlation between WM and Gf to range from 0.40 to 0.80, with many correlations being closer to the lower end of this range (Colom, Flores-Mendoza, & Rebollo, 2003; Fry & Hale, 1996; Jurden, 1995; Salthouse, Mitchell, Skovronek, & Babcock, 1989; Tucker & Warr, 1996; Verguts & De Boeck, 2001). These correlations can differ as a function of the statistical analyses performed, the number and type of tests used to measure the two constructs, the sample size, and the fact that many of these studies used a student sample, among other things (Gignac, 2014; Oberauer, Schulze, Wilhelm, & Süß, 2005; Yuan, Steedle, Shavelson, Alonzo, & Oppezo, 2006). Nonetheless, most WM researchers generally conclude that the relationship between WM and fluid reasoning is relatively strong but the conclusion that WM and Gf are isomorphic is not warranted.

Theoretically, it is not surprising that WM and Gf have moderate correlations. After all, the entire foundation of psychometric “g” is posited on the positive manifold among cognitive tests. Therefore, all lower-level cognitive abilities are moderately correlated with “g” or Gf to some extent (Carroll, 1993). However, based on abundant literature suggesting more moderate correlations as well as the cognitive processes associated with WM, WM scores may have predictive validity above and beyond measures of “g”.

WM and Learning

Alloway and Alloway (2010) set out to determine whether WM simply served as a proxy for IQ when predicting academic attainment. The researchers administered WM measures and IQ

measures to 98 kindergarten students with a mean age of five. Six years later, these same students were given two tests of academic attainment and learning ability. Their results showed that not only did WM predict a significant amount of variance in learning outcomes, but it also accounted for more variance than IQ. They concluded that WM is a cognitive skill separate from intelligence and important in understanding academic success.

Luo, Thompson, and Detterman (2006) had similar findings in their study with a slightly older sample. Among children and adolescents between 6 and 19 years of age, they compared the predictive validity of a battery of WM tests and a battery of Gf tests. After taking into account the predictive power of the Gf tests, the WM tests predicted an additional 11-14% variance in scholastic achievement scores. They concluded that a WM battery should be utilized in further studies when predicting learning and educational outcomes.

Furthermore, Krumm, Ziegler, and Buehner (2008) used undergraduate students to examine the incremental validity explained by WM measures above fluid reasoning measures in predicting German students' language and science school grades. Although no significant variance was explained by WM measures for science grades, WM measures added significant predictive power for language grades above and beyond measured of reasoning ($p < .01$).

De Jong and Das-Smaal (1995) reported findings which also strengthen the case for WM as having significant incremental prediction over "g". They found that when predicting scholastic achievement test scores in over 2,000 fourth-graders, WM measures added incremental validity over "g".

Past literature has made a convincing case that WM could be important in providing extra predictive power in the prediction of learning and educational outcomes above "g" or Gf.

To supplement this literature, WM has also been shown to have high correlations with other constructs that may be linked to learning such as following directions (Engle et al. 1991), language comprehension (King and Just 1991), note-taking (Kiewra & Bentley, 1991), computer-language learning (Kylönnen & Stephens, 1990), and multi-tasking (Hambrick et al., 2010).

Since the evidence has been presented for the effect of WM on learning, it is logical to extend this by examining the effects WM may have on task performance, as the acquisition of task knowledge is important for task performance.

WM and Task Performance

Theoretically, it is intuitive that one's ability to maintain and manipulate newly acquired information while blocking out distracting information can predict task performance. When a person is able to hold ideas or representations in accessible cognition while performing other tasks, then revisiting the information without sacrificing the integrity of that information, (s)he is probably more likely to perform better on the task.

Not much has been done using WM measures to predict task performance of employees (Kane et al., 2007). Two studies of note are Bosco et al.'s (2015) studies in which WM (which they called executive attention) measures were compared to the WPT in predicting simulation performance of undergraduate students and supervisor-rated job performance of bank employees. They found WM to have a similar or better predictive validity as the WPT in both studies. The WM measures had validity coefficients of 0.22 and 0.21 for simulation performance and supervisor-rated performance, respectively, in the two studies. Although these results are promising for the WM case, more needs to be done to support the validity of WM in predicting performance.

Military organizations have also examined the role of WM in predicting task performance. Wolfe et al. (1995) found that a WM measure predicted unique variance above the ASVAB in task performance in one-third of Navy technical schools. The most unique variance accounted for by the WM measure was 5%. This is consistent with Held and Wolfe's (1997) finding; a WM measure predicted 5-10% additional variance in task performance related to basic Air Force training modules, advanced Air Force training modules, and Operations Specialist ratings.

Thus, there is limited literature on how WM may be able to incrementally predict these important outcomes. Providing additional evidence for the incremental predictive validity of WM is important in that it gives us a better understanding of how WM relates to these constructs. Another way to better understand WM is to more closely examine the functions of WM, as well as how these functions contribute to learning and task performance outcomes.

Working Memory: Capacity and Processing

When Baddeley and Hitch (1974) first proposed their WM model, they conducted ten experiments that required participants to perform different cognitive tasks while maintaining information in memory to be recalled after task completion. They found a trade-off between the amount of memory storage required (i.e., the number of pieces of information to be remembered) and the participants' performance on the cognitive tasks, such that more information to be remembered led to poorer performance on the simultaneous cognitive tasks.

Since then, WM has been considered as having two closely related functions: capacity and processing (Baddeley & Hitch, 1974; Baddeley, 1986). WM capacity (WMC) is the amount of information that can be held in WM. WM processing (WMP) is the ability to successfully maintain, update, and manipulate this available information while allocating cognitive resources

toward a goal (Kane et al., 2007). Most WM researchers conceptualize the WM system to consist of both WMC and WMP (Daneman & Carpenter, 1980; Kyllonen & Christal, 1990; Oberauer, Süß, Wilhelm, & Wittmann, 2003; Salthouse, 1991).

From an individual differences perspective, Daneman and Carpenter (1980) reinforce the dual function model. They posit that people with better WM systems are better readers because of the interplay of WMC and WMP; an individual with greater WMC will process text more efficiently through parsing, integrating and decoding of information. As such, the relationship between WMC and WMP is dynamic in that greater WMC increases processing efficiency and increased WMP efficiency in turn frees up WMC (Daneman & Carpenter, 1980). Thus, an individual with a greater WMC will also be advantaged in WMP capabilities.

This dual-function model of WM provides the logical underpinning for the hypothesized model whereby knowledge only partially mediates the WM-task performance relationship (see Figure 1). That is, both WMC and WMP use cognitive processes which enhance the ability of individuals to learn new material. WMC and WMP function together in the learning of new material such that individuals are not only able to store new information they are presented with, but they are also able to maintain, manipulate, and update this information without sacrificing the integrity of this information (Kane et al., 2007). Thus, both WMC and WMP are responsible for the acquisition of task knowledge.

I propose that WMP will also enhance task performance directly. WMP is important in being able to block out distracting stimuli, update stored information, and manipulate information in the WM system; these are essential processes for performing cognitive tasks. Also, WMP allows individuals to engage in these cognitive processes while utilizing other cognitive resources toward reaching a goal. Although the storage of new information (i.e.,

WMC) may not be as important for task performance, I suggest that WMP will drive the direct effect of WM on task performance.

To illustrate this in a real-world scenario, consider two waiters working at a restaurant. Both waiters have been working in their positions for the past three years. One waiter is slightly more intelligent than the other, so it is safe to say based on the research (e.g., Schmidt & Hunter, 2004) that this waiter probably learned the relevant job knowledge faster than the less intelligent waiter. This job knowledge could be learning the menu items, how to arrange the tables, the best way to take orders, how long to wait between the appetizers and meals, and the list goes on. However, over the past three years, the less intelligent waiter has learned all of these things as well, to the point where both waiters know an equal amount of job knowledge, and could probably tell you anything you would need to know about their current position.

However, consider that the less intelligent waiter has a better WM than the more intelligent waiter. During the rush hour, both waiters are flooded with diners and they are forced to keep a constant representation in their heads consisting of where each of their tables is in the process of ordering drinks, appetizers, entrees, and desserts. Furthermore, they have the relevant task knowledge such as how long they are supposed to wait before asking the table if they would like an appetizer. Keeping track of the dining timeline of different tables, listening to customer complaints, making sure the orders are correctly placed, making sure the busboys are cleaning the tables, and interacting with the customers to ensure great customer service all requires cognitive processes associated with WM. Therefore, when both waiters eventually reach equal levels of job knowledge, the waiter with the better WM will have the advantage in performing the job tasks effectively because of his better WMP comparative to the waiter with the worse WMP. This exemplifies the partially mediated model presented in Figure 1.

In conclusion, the present study had two focal purposes. It attempted to support the notion that WM is able to predict task performance through the mediation of task knowledge. Further, it aimed to empirically support task knowledge as a partial mediator of the WM – task performance relationship.

The Current Study

The current study aimed to better understand WM and how it relates to three important constructs: general intelligence, task knowledge, and task performance. Limited research has assessed the incremental validity of WM in predicting learning and task performance above “g”. However, this is probably because in the literature, other specific abilities of “g” have been found to be inadequate in accounting for additional variance in these outcomes. Also, measures of “g” and Gf are so established, convenient, and cost-efficient that many researchers have stopped trying to improve upon the predictive validity of these measures. However, the proposed study posits that important cognitive processes are involved in WM, and these processes are not measured by typical measures of “g” or Gf.

As stated previously, the first purpose of this study was to provide evidence that WM scores can predict learning and task performance above measures of “g” or Gf. The present study theorizes that WM allows people to be better able to learn new information they are presented with. That is, people who are able to hold more relevant information in WM and attend to the details of the new information and manipulate this information while inhibiting irrelevant stimuli will be more likely to retain the information. That being said, this increase in the learning of task knowledge will, in turn, allow them to be better able to perform tasks which apply this knowledge.

However, if WM predicts the same variance in learning and task performance that Gf predicts, then WM measures would exhibit the same predictive validity as Gf. The proposed study posits that WM scores will add incremental predictive validity above Gf, predicting variance in learning and task performance that Gf does not predict. Indeed, the RAPM, Wonderlic, and other commonly used tests of Gf do not include measures that effectively assess a person's ability to store and manipulate information while blocking out distracting information. Lastly, the cognitive processes that WM is involved in seem as relevant to successfully carrying out tasks as those of fluid reasoning captured by more common measures of cognitive ability. Thus, the current study hypothesizes:

Hypothesis 1: After controlling for fluid intelligence (Gf), WM will account for unique variance in task knowledge.

Hypothesis 2: After controlling for fluid intelligence (Gf), WM will account for unique variance in task performance.

As Daneman and Carpenter (1980) point out, people with better WM benefit because of the dual functionality of the WM system (i.e., WM serving as capacity and processing). People with greater WMC are able to store more information for longer periods of time in the presence of distracting or irrelevant stimuli (Kane et al., 2007) and people with better WMP are better able to attend to, store, maintain, update, and manipulate information as it is presented. Furthermore, benefits in one functionality results in benefits of the other functionality, such that people with greater WMC have better WMP and vice versa.

Therefore, the present study suggests that because WM has both WMC and WMP functions, WM scores will predict acquisition of task knowledge because of the importance of

both of these cognitive functions for learning. This study takes this perspective one step further, and suggest that this learning of task knowledge will only partially mediate the relationship between WM and task performance. This mediation is partial because of the importance of the WMP function during task performance. That is, WMP's cognitive processes of inhibition, manipulation, and updating will be important for the actual performance of any cognitive task, independent of the task knowledge that is acquired. Thus, a partially mediated model is hypothesized (Figure 1):

Hypothesis 3: Controlling for general intelligence, task knowledge will partially mediate the relationship between WM and task performance.

The current study tested these hypotheses by administering WM and “g” tests and examining participants’ (undergraduate students’) ability to learn and perform a learning-related task. Participants were presented with a video tutorial providing them with details of how to properly format an APA reference page. APA knowledge pre- and post-tests will assess knowledge gains from the tutorial. Participants will then be asked to apply this knowledge by creating their own APA reference page. Thus, the proposed study will examine the predictive validity of both WM and “g” in the prediction of both task knowledge and task performance outcomes.

Chapter 3

Methods

I. Participants

Participants were 167 undergraduate students recruited from the Psychology Department of a Southeastern university. Of the final 154 participants, the demographic breakdown was 15% Asian-American, 10% African-American, 7% Hispanic, 68% White, and 65% female.

Participants were recruited by posting flyers, making announcements in Psychology class lectures, and posting the study information on a Psychological research website that is accessible to all undergraduate Psychology majors. Participants received extra credit in one of their Psychology courses for participation in the study; participants were also entered into a raffle with the prizes being a \$50 gift card and a \$20 gift card.

II. Procedure

Participants signed up for the laboratory study via an online psychology experiment management system. Participants were brought into a computer lab in groups of between one and

six. After signing consent forms, participants were assigned to their own computer and completed pre-task assessments of general aptitude and WM. After the Gf and WM tasks, participants were given a proactive break where they were asked to close their eyes and think of a calm, relaxing environment for five minutes at their computers. This was meant to reduce fatigue. Participants were then given a pre-test to assess their knowledge of constructing reference pages in APA style. Participants were then instructed to view an APA-produced video tutorial on constructing references in APA style and they were told their task will be to construct a reference list in accordance with APA style guidelines. After viewing the video, participants were given another knowledge test of APA style guidelines, after which they were given the material that they will use to construct a reference list.

Experimental Task. After the initial WM, g, and APA knowledge assessments were completed, participants were given a nine-minute, thirty-one second video tutorial which instructed them on how to properly format a reference page given the information of a book or journal article. This tutorial served as an opportunity for them to learn material which most of them had relatively little knowledge about. The tutorial described in detail, with examples, the correct formatting needed to create a reference page. The video can be found at <https://www.youtube.com/watch?v=XNGHw6Gwt7Q>.

Following the video, participants were given a sheet of paper with the necessary information about a set of journal articles (Appendix A) for which they were asked to create a properly-formatted reference page using Microsoft Word. Each journal article they were given consisted of author name(s), year published, article title, journal title, journal volume number, and page numbers. Reference pages were scored based on their order of references, punctuation

(e.g., commas and periods in the right place), parentheses, capitalization, spacing, italics, spelling, and order of items in the reference.

Based on the idea that many students may not know how to create hanging indents and may not know if font type or font size are relevant for the task, a pre-created Word document template was already on the desktop of each computer. These templates were already set to Times New Roman, 12-pt font, with hanging indents formatted.

Participants were allotted 12 minutes to compile the reference page, at the end of which they were asked to stop typing and save the document. Participants were thanked for their participation, and their Word documents were printed for scoring. All of the measures were taken in a computer lab, and were completed in one 90-minute session.

IV. Measures

APA-style knowledge tests. Two forms of knowledge test were developed (See Appendix B) based on rules and guidelines for proper APA formatting (<https://owl.english.purdue.edu/owl/resource/583/03/>). Each form consisted of ten multiple-choice questions, with each question having four or five response options. Participants were given eight minutes to complete the knowledge test both at pre-test and post-test. These tests scored based on number of correct responses.

Automated OSPAN. Participants started off with on-screen instructions explaining the OSPAN assessment task, followed by three practice blocks. The purpose of the practice task was two-fold; first, participants were afforded the opportunity to practice with the task, and second, response latency was used to establish the time that each stimulus is presented. The basic OSPAN task requires an individual to remember and report individual letters that are given to them between answering simple math problems. A math problem [e.g., $(4 \times 2) - 3$] is presented

on the screen for a short period of time, determined by response latencies collected during the practice block. Immediately following this math problem the screen displays a proposed solution to the math problem (e.g., “5”), where individuals must indicate whether the presented solution is correct by indicating “True” or “False” by clicking either word on the screen. Within the instructions and practice problems, participants were encouraged to get at least 85% of the items correct, or else the data would not be recorded for them.

Within each block of problems, after the participant chose between true and false, a screen containing one individual letter appears for a short time period determined by response latencies collected during the practice block. Following the presentation of the letter, sequence is repeated with a new problem and new letter. Each block contains between three to seven math problems and letters to be remembered. Block size is not constant so that participants cannot predict the amount of numbers they are supposed to remember, challenging both the processing and capacity functions of WM. This block variation was consistent with past research using the RSPAN and OSPAN (Salthouse & Pink, 2008; Unsworth, Heitz, Schrock, & Engle, 2005).

There were ten total blocks for the OSPAN, with each size block (i.e., 3, 4, 5, 6, 7) appearing twice. At the end of each block, participants were given a recall screen, where they were instructed to report the letters presented in the most recent block. Participants were instructed to report them in chronological order, and were told to report a “BLANK” to hold the spot of a letter which they did not remember. No time limit was given for the recall screen. Once each participant submitted his/her answers, a new block of math problems began.

Scoring was based on how many letters the participant accurately remembered in the correct place on the recall screen. Therefore, the highest possible score was 50 (i.e., ten blocks with an average of 5 letters in each block). This scoring technique is called partial-credit load

scoring and the OSPAN scores have been shown to have an internal consistency of .808 (Conway, Kane, Bunting, Hambrick, & Wilhelm, 2005) when using this technique.

Participants were given feedback throughout the task on their percent correct for the math problems, and were encouraged to get at least 85% of the math problems correct for the data to be recorded. Although data on the accuracy of math solutions were not used in the analyses, the accuracy reflects how invested individuals are in solving both tasks accurately rather than just memorizing the letters. Per Conway, Kane, Bunting, Hambrick, and Wilhelm's (2005) recommendations, data from a participant were not used if he/she scored less than 85% correct on the math problems. The average time to complete the OSPAN was 12.3 minutes.

Automated RSPAN. The Automated Reading Span Task (RSPAN) was virtually identical to the OSPAN task, with one difference. Instead of using the math problems as distracting information, the RSPAN presents a sentence such as "Because she gets to knife early, Amy usually gets a good parking spot." Participants then indicated if a sentence makes sense by indicating "True" or "False" by clicking either one of them presented on the screen. Although similar, the RSPAN and OSPAN differ in the secondary processing (i.e., interfering) tasks they present; the OSPAN introduces quantitative processing, while the RSPAN introduces verbal processing. Thus, both were used to capture both types of processing for secondary tasks.

Just as in the OSPAN, participants were given practice items first, and a baseline time was calculated which determined the amount of time participants were given to choose an answer. Participants were then presented with a letter to be recalled, just like the OSPAN. The ten blocks contained three to seven letters to be remembered, totaling a score of 0-50 correct letters in the correct position.

The RSPAN task and OSPAN task were separated by the Progressive Matrices Task (PMT) in order to reduce perceived boredom on behalf of the participants. The OSPAN and RSPAN were counterbalanced so participants were randomly assigned to the RSPAN or OSPAN task first, followed by the PMT, and then given the remaining WM task. Similar to the RSPAN, data were not used from any participant failing to achieve 85% correct on whether the sentence was logical. Internal consistency of the RSPAN scores using the partial-credit load scoring has been reported at 0.776 (Conway et al., 2005). Screenshots of the RSPAN and OSPAN can be found in Appendix C. The average time to complete the RSPAN was 11.5 minutes.

The WM tasks were all administered using Inquisit 5 software by Millisecond. The Inquisit software allows researchers to administer established cognitive laboratory measures as well as customize these measures. The software allows for a baseline choice reaction time to be calculated for each participant, and to apply these baselines for the remainder of the tasks.

Gf. The progressive matrices task (PMT) is a readily accessible online task developed and used by Mensa® to measure IQ and is inspired by Raven's Advanced Progressive Matrices (RAPM). The RAPM has shown to be a reliable and valid measure of fluid reasoning (Gf), and it is generally recognized that "g" is highly saturated with Gf (Carpenter, Just, & Shell, 1990; Spearman, 1946; Vernon & Parry, 1949), and Gf is considered to have the highest g-loading compared to other cognitive ability tests. The PMT presents participants with a matrix of boxes filled with shapes and patterns, with one box missing. The participants must determine the rule that explains the shapes or patterns, and use this rule to determine which of the eight options provided best completes the pattern. The PMT contained 20 matrices, starting with the easiest items, and progressively reaching the most difficult items. Participants were given 20 minutes to complete this task. Scores were based on number of correct responses out of the 20 items.

Task performance. Participants were given 12 minutes to create the reference page. In total, the participants could score a total of 41 points. Participants could earn eight points for each of the five references they are tasked to cite. One point was given for each of the following criteria: 1) spelling everything correctly throughout the reference (e.g., authors' names, titles), 2) inserting commas in the correct places throughout the reference, 3) inserting periods in the correct places throughout the reference, 4) inserting periods in the correct places throughout the reference, 5) proper capitalization throughout the reference, 6) proper spacing throughout the reference, 7) proper italics throughout the reference, and 8) placing each item of the reference in the proper order. A final point was awarded for putting the five references in the proper order on the reference page, for a total of 41 points. The scoring rubric for the APA reference page can be found in Appendix D.

VI. Analysis

Pre-test scores and scores on the PMT were statistically controlled for by regressing post-test scores on pre-test scores and PMT scores, hierarchically. Pre-test scores were entered first as they are more proximal to the post-test scores. The deleted residuals were saved so that the new post-test scores had pre-test and PMT scores parsed out. The same procedure was followed for the APA task scores.

A non-parametric bootstrap was completed using the “mediate” function in R in order to examine both indirect and direct effects of WM on task performance. Bootstrapping is recommended over other tests for mediation such as the Sobel test because of the Sobel test's lack of power. Further, bootstrapping is not reliant on a normal sampling distribution (Hayes, 2013). Bootstrapping involves resampling with replacement, where effects are computed from a random sample of the data, and this is replicated over many simulations. From these simulations,

a sampling distribution can be empirically approximated, and a confidence interval can be determined for both the direct and indirect effects. Specifically, percentile bootstrapping was conducted rather than bias-corrected bootstrapping because bias-corrected bootstrapping tends to be too liberal, with alpha approximately 0.07 (Fritz, Taylor, & MacKinnon, 2012). Percentile bootstrapping is recommended to reduce Type I error rate (Hayes & Scharkow, 2013). Ten thousand simulations were completed, and 95% confidence intervals were computed for both effects as well as total effects.

Chapter 4

Results

Descriptives and Correlations

Out of the sample of one-hundred sixty-seven, two outliers were removed because they scored greater than four standard deviations below the mean on the post-test. Participants who scored eight or higher on the pre-test were also taken out of the analysis to mitigate ceiling effects for task knowledge. This resulted in a final sample of $n = 154$. Lastly, there were two items on the measure of task knowledge that were taken out of the analyses due to the fact that 98% and 99% of participants answered these questions correctly, resulting in an 8-item measure of task knowledge. Cronbach's alpha for this measure was 0.63. Cronbach's alpha for the 10-item pre-test was 0.59. The descriptive statistics for the pre-test, Gf, RSPAN, OSPAN, task knowledge, and task performance appear in Table 1.

Table 1.
Descriptive Statistics

	Minimum	Maximum	Mean	Std. Deviation
Pre-test	1	7	4.84	1.36

Gf	5	19	12.69	2.46
OSPAN	14	50	39.56	8.64
RSPAN	13	50	36.75	7.39
TK	1	8	6.07	1.25
TP	7	41	28.06	8.87

Notes. Pre-test = Task knowledge before tutorial. OSPAN = Operation Span. RSPAN = Reading Span. TK = Task knowledge after tutorial. TP = Task performance. $n = 154$.

Examining gender, pre-test scores, and “g”, preliminary analyses revealed no interaction effects for either task knowledge or task performance. Furthermore, age, year in school, major, and gender were not related to either task knowledge or task performance. There was a race effect for “g”; Asian/ Pacific Islanders ($n = 23$) scored higher on “g” than African-Americans ($n = 16$; $t = 2.38$, $p < .05$). It can sometimes be an issue that using Asian-Americans who speak English as a second language to complete a verbally loaded task may hinder performance on those tasks. However, Asian-Americans scored slightly higher than other groups on both WM and Gf tasks, so this does not seem to be an issue even though English as a second language was not measured. The bivariate correlations for the pre-test, “g”, OSPAN, RSPAN, task knowledge, and task performance appear in Table 2.

Table 2
Correlation matrix of measured variables.

	1	2	3	4	5	6	7
1. Pre-test	—						
2. Gf	.05	—					
3. OSPAN	.16*	.22**	—				
4. RSPAN	.18*	.19*	.46**	—			
5. WM	.20*	.24**	.88**	.83**	—		
6. TK	.04	.18*	.19**	.18*	.22**	—	
7. TP	.21**	.16*	.14*	.30**	.24**	.48**	—

Notes. Pre-test = Task knowledge before tutorial. OSPAN = Operation Span. RSPAN = Reading Span. TK = Task knowledge after tutorial. TP = Task performance. WM = WM composite. $n = 154$. $**p < .01$, $*p < .05$ (one-tailed)

From Table 2, it is evident that the RSPAN-OSPAN correlation ($r = .46$) is more moderate than the correlation typically found between WM measures. This can sometimes happen as a result of the importance the researcher places on the secondary task. Participants who invest much attention into performing well on the secondary tasks may have different scores as a result of the student being more/less efficient at processing verbal vs. quantitative information. The correlation may have been larger if the participants only needed to get 50% correct on the secondary task versus 85%, as the secondary task would have played less of a part in the WM scores.

Because of the relatively modest correlation ($r = 0.46$) between the WM measures (i.e., RSPAN and OSPAN), the two measures were not combined to form an overall WM composite score. Rather, the two WM measures were tested separately resulting in two sets of analyses.

Residualizing Outcome Variables for Mediation Tests

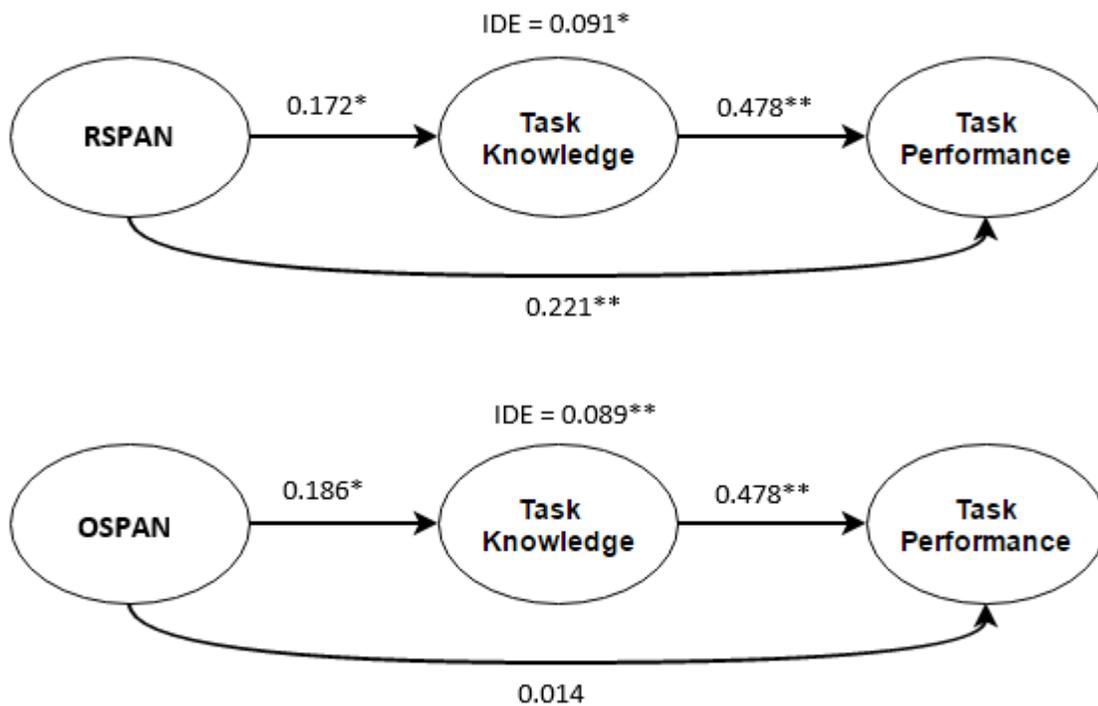
In order to control pre-test knowledge for the mediation analyses, post-training knowledge and task performance were regressed on pre-test scores. For both post-training task knowledge and task performance, the deleted residuals from these regression analyses were used in the mediational analyses. Deleted residuals are calculated by fitting a regression model with each observation deleted, one at a time, from the data. The predicted value based on the regression model without the observation is subtracted from the observed value of the observation, and the difference between the two is the deleted residual. Thus, these residual values represent the outcome variables with pre-test knowledge statistically controlled.

When creating these residuals, it was found that pre-test knowledge ($\beta = 0.037$; $p = 0.324$) did not affect task knowledge; pre-test scores accounted for 0.1% of variance in task knowledge. Pre-test scores ($\beta = 0.212$; $p < .01$) predicted task performance; pre-test scores accounted for 4.5% of variance in task performance.

Baseline Mediation Model Not Controlling for Gf

Baseline mediation analyses were conducted without controlling for Gf. Nonparametric bootstrapping mediation with ten thousand simulations was conducted using the RSPAN and OSPAN as mediators separately. Results for both path models appear in Figure 5.

Figure 5. Beta coefficients for the model using the RSPAN and OSPAN, not controlling for Gf.



Notes. IDE = Indirect effect. **Beta coefficients significant at the $p = 0.01$ level (one-tailed). *Beta coefficients significant at the $p = 0.05$ level (one-tailed).

The RSPAN, accounted for 2.9% of variance in task knowledge ($\beta = .172$; $p < .05$). The RSPAN scores had an indirect effect of 0.091 on task performance, as mediated by task knowledge ($p < .05$). Further, RSPAN scores had a direct effect of 0.221 on task performance ($p < .01$). The total effect of RSPAN on task performance was 0.312 ($p < .01$) and the RSPAN accounted for 6.9% of total variance in task performance. Therefore, when running mediation analyses without controlling for Gf, the partial mediation model was supported using RSPAN scores.

The OSPAN, accounted for 3.5% of variance in task knowledge ($\beta = .186$; $p < .05$). The OSPAN scores had a significant indirect effect on task performance, as mediated by task knowledge ($\beta = 0.089$; $p = .01$). OSPAN scores did not have a direct effect on task performance ($\beta = 0.014$; $p = .44$). The total effect of OSPAN on task performance was 0.104 ($p = .12$). Therefore, when running mediation analyses without controlling for Gf, the partial mediation model was not fully supported using OSPAN scores, but rather there was full mediation through task knowledge. In sum, without controlling for Gf, there is a partially mediated effect of RSPAN scores on task performance, and a fully mediated effect of OSPAN scores on task performance through task knowledge.

Mediation Model Controlling for Gf

Mediation analyses were then conducted while controlling for both pre-test knowledge and Gf to determine if the effect of WM on task performance holds even after taking Gf into account. In order to control pre-test knowledge and Gf for mediation analyses, deleted residuals were once again saved after regressing task knowledge and task performance on both pre-test scores and Gf. Pre-test scores were entered in the first step, followed by Gf. These new deleted

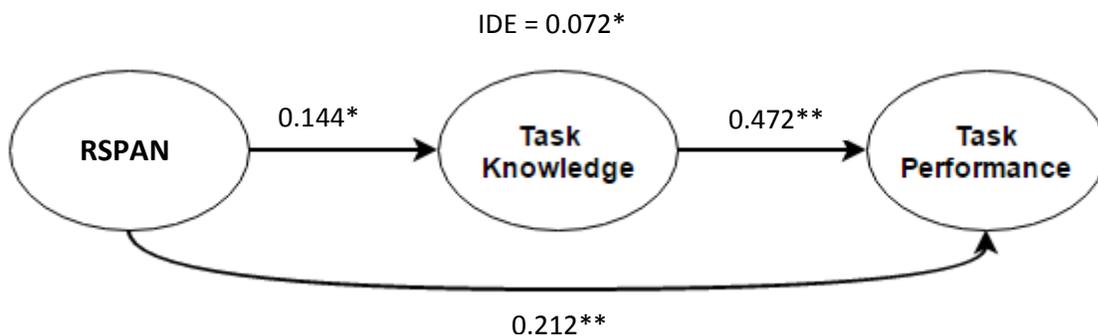
residuals represent task knowledge and task performance after statistically controlling for both pre-test scores and Gf.

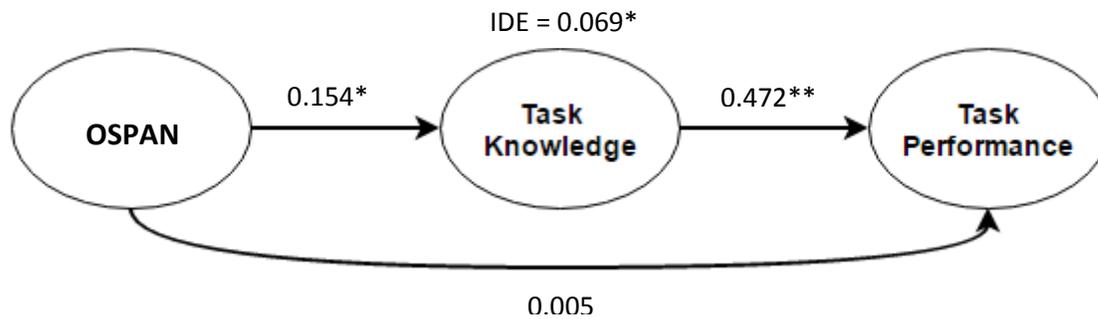
When creating these residuals, it was found that pre-test knowledge ($\beta = 0.039$; $p = 0.312$) did not affect task knowledge, but Gf ($\beta = 0.184$; $p < .05$) did. Together they accounted for 3.6% ($p < .05$) of variance in task knowledge. Pre-test knowledge ($\beta = 0.204$; $p < .01$) and Gf ($\beta = 0.143$; $p < .05$) both affected task performance. Together they accounted for 6.6% ($p < .01$) of variance in task performance.

In order to determine the amount of incremental variance explained by WM and to better interpret Hypotheses 1 and 2, deleted residuals controlling just pre-test scores were created by regressing task knowledge and task performance on pre-test scores. This way, Gf could be entered in the first step of the regression analyses, followed by WM in order to determine amount of incremental variance explained by WM.

The mediation model and path coefficients can be found in Figure 6. The models show that even after controlling for Gf, the mediation results do not change – task knowledge partially mediates the relationship between RSPAN and task performance, and task knowledge fully mediates the relationship between OSPAN and task performance. Each hypothesis is interpreted individually below.

Figure 6. Beta coefficients for the model using the RSPAN and OSPAN, controlling for Gf.





Notes. IDE = Indirect effect. **Beta coefficients significant at the $p = 0.01$ level (one-tailed). *Beta coefficients significant at the $p = 0.05$ level (one-tailed).

Hypothesis 1. Hypothesis 1 was tested by examining the effect of both the RSPAN and the OSPAN on task knowledge, separately, above and beyond Gf. Gf explained 3.2% of the variance in task knowledge ($\beta = .180$; $p < .05$). RSPAN explained an additional 2.1% variance ($\beta = .144$; $p < .05$) and OSPAN explained an additional 2.4% variance ($\beta = .154$; $p < .05$) in task knowledge. Therefore, Hypothesis 1 was supported using both measures of WM.

Hypothesis 2. Hypothesis 2 predicted WM accounts for unique variance in task performance. Gf explained 2.2% of the variance in task performance ($\beta = .147$; $p < .05$). RSPAN explained an additional 5.7% variance ($\beta = .238$; $p < .01$) and OSPAN explained an additional 0.5% variance ($\beta = .074$; $p = 0.186$) in task performance. Therefore, hypothesis 2 was supported using RSPAN scores but not supported when using OSPAN scores.

Hypothesis 3. Hypothesis 3 predicted task knowledge partially mediates the relationship between WM and task performance. For the RSPAN scores the direct effect of WM on task performance remained ($\beta = .212$; $p = 0.01$) remained after controlling for Gf. Although the indirect effect of RSPAN scores failed to reach the $p = 0.05$ significance threshold ($\beta = 0.072$; p

= 0.055), it was concluded that hypothesis 3 was supported and RSPAN had both a direct and indirect effect on task performance.

The OSPAN scores partially mediated model found no support for the direct effect of OSPAN scores on task performance ($\beta = .005$; $p = 0.48$), but found an indirect effect on task performance through task knowledge ($\beta = .069$; $p < .05$). Thus, the hypothesis that task knowledge partially mediates the relationship between WM and task performance when controlling for Gf was not fully supported when using the OSPAN scores. Rather, the full mediation of the OSPAN-task performance relationship holds after controlling for Gf.

Exploring Both WM and Gf as Antecedents

Exploratory regression analyses were conducted to examine the results when WM and Gf were both entered as antecedents. Task knowledge and task performance were regressed on Gf and the WM measures (separately) at the same step.

For RSPAN, both Gf ($\beta = 0.154$; $p < .05$) and the RSPAN scores ($\beta = 0.144$; $p < .05$) affected task knowledge, accounting for 5.2% ($p < .01$) of the variance in task knowledge. The RSPAN ($\beta = 0.243$; $p < .01$) affected task performance, but Gf ($\beta = 0.102$; $p = 0.102$) did not, while together they accounted for 7.9% ($p < .01$) of the variance in task performance.

For OSPAN, both Gf ($\beta = 0.147$; $p < .05$) and the OSPAN scores ($\beta = 0.154$; $p < .05$) predicted task knowledge; together they accounted for 5.5% ($p < .01$) of the variance in task knowledge. Neither Gf ($\beta = 0.131$; $p = 0.113$) nor the OSPAN ($\beta = 0.074$; $p = 0.186$) predicted task performance, while together they accounted for 2.7% ($p = 0.065$) of the variance in task performance.

Summary

In summary, the two WM measures were kept separate in the mediation analyses due to their moderate correlations. Without controlling for Gf, a partially mediated model was supported using the RSPAN. Using the OSPAN, there is evidence for full mediation through task knowledge. When controlling for Gf, these results remain. Hypothesis 1 was supported by both measures of WM, as both measures predicted task knowledge above Gf. Hypothesis 2 was supported using the RSPAN scores as RSPAN scores significantly predicted task performance above Gf. However, hypothesis 2 was not supported using the OSPAN scores. The RSPAN-task performance relationship was partially mediated by task knowledge, and hypothesis 3 was supported using the RSPAN. Although there was an indirect effect found for the OSPAN, there was no direct effect found for the OSPAN. This provides evidence for a fully mediated model using the OSPAN.

Chapter 5

Discussion

The current study examined the effects of task knowledge as a partial mediator when controlling for Gf. This mediated model adopted an established model in which task knowledge mediates the relationship between Gf and task performance. This study has findings that can begin to measure underlying processes that affect both these outcomes. When controlling for “g”, the hypothesized partial mediation model was supported using the RSPAN scores, but task knowledge fully mediated WM and task performance for OSPAN scores.

Conceptual Implications

Support for the WM construct is important because, unlike g, WM provides processing explanations as to both the acquisition of task knowledge and variability in task performance. Although researchers recognize that psychometric “g” is a ubiquitous predictor of outcomes (Jensen, 2002), there is no processing explanations for the effects of “g” on outcomes. As described earlier, psychometric “g” is mostly considered an artifact resulting from the positive manifold of cognitive tests rather than a meaningful latent construct (Jensen, 1986). Much of the

literature extends this to say that individuals who do well on these cognitive tests (i.e., have a higher score on “g”) would naturally experience better outcomes. However, research regarding the causal processes leading to these outcomes is scant. As Jensen (2002) explains, “...the causal connections in the whole nexus of the many diverse phenomena involving the g factor is highly complex” (p.39).

The current study focused on a small piece of this complex puzzle by examining two causal paths by which cognitive abilities may affect outcomes. Specifically, this study theorized that the WMC function of WM enables individuals to retain task knowledge by blocking out distracting information while learning, storing, and rehearsing newly presented information. At a higher level of abstraction, it is well known that task knowledge is a causal antecedent of task performance (e.g., Palumbo, Miller, Shalin, & Steele-Johnson, 2005; Schmidt, Hunter, & Outerbridge, 1986). Further, the WMP function of WM serves as a means of being able to maintain, update, and rehearse newly learned information while successfully processing another task and inhibiting stimuli that would inhibit goal achievement. These two functions, as measured by the two WM measures in the current study, explain processes by which cognitive aptitude can affect outcomes. Therefore, not only does WM provide prediction, but it also furthers our understanding of theory and the mechanisms by which learning and performance occur.

As Karl Popper famously said, “A theory that explains everything, explains nothing.” This quote is relevant for the research literature of examining the broad theory of “g”. Most researchers conclude that “g” leads to a wide range of better outcomes, but do not theorize the specific processes which would help to explain relevant mechanisms which contribute to these outcomes. By exploring these lower level processes associated with specific abilities such as

WM, researchers can better understand how cognition and processing contribute to learning and task performance; this achieves a more specific understanding of mental abilities that produces falsifiable hypotheses.

Regarding the results of the current study, the low correlation between “g” and WM ($r = 0.22$ and 0.19 for OSPAN and RSPAN, respectively). Researchers typically find that the correlation between “g” and measures of WM ranges from 0.40 and 0.80 (Colom, Flores-Mendoza, & Rebollo, 2003; Fry & Hale, 1996; Jurden, 1995; Salthouse, Mitchell, Skovronek, & Babcock, 1989). It could be that because these correlations are weaker than usual, “g” is not properly getting captured, and therefore, WM measures are easily able to add incremental validity to the measure of “g”. However, in the current study, the criterion validity correlations for “g” were 0.18 and 0.16 for task knowledge and task performance, respectively, which are similar to criterion-related validity coefficients for “g” in other contexts. Furthermore, the theory in this study posits that WM tasks measure underlying processes (i.e., updating, rehearsing, maintaining, inhibiting) that are: 1) important for learning and performance, and 2) not represented by measures of “g” typically used (e.g., the RAPM). It is because of these processes that WM affects task knowledge and task performance beyond “g”.

This raises the question as to whether WM should be considered a specific ability of “g” which should be included in a general cognitive ability test battery, or if WM should be considered and measured separately from “g”. Theoretically speaking, psychometric “g” is based on the positive manifold between cognitive ability tests, and therefore measures of specific abilities such as WM contribute to the overall hypothetical “g”. But in a more practical sense, as mentioned above, WM measures specific unique lower-level processes that aren’t reflected when practitioners measure general cognitive ability. Although Gf (e.g., RAPM) and Gc (e.g.,

ASVAB) tests are widely used to measure overarching “g”, they measure fluid reasoning and comprehension knowledge processes, respectively. By definition, WM falls under the hierarchy of “g”; however, most measures of “g” do not include measures of WM. As such separate measures of WM should be considered for inclusion when studying mental ability.

In the current study, the correlation between the RSPAN scores and OSPAN scores was relatively modest ($r = 0.46$); therefore, separate path models were tested. This correlation is smaller than many of those reported in the literature for these measures, such as 0.62 (McVay & Kane, 2012), 0.67 (Bailey, Dunlosky, & Kane, 2008), and 0.71 (Unsworth, Heitz, Schrock, & Engle, 2005). Although both the RSPAN and OSPAN were meant to represent the WM latent construct given their record of larger correlations, this finding may have implications for the domain-specificity of WM.

Domain Generality vs. Specificity of WM

Researchers ascribing to WM as domain specific posit that people have different WM capabilities as a function of the type of information being processed. In the context of this study, these domains are verbal and quantitative information, and when adopting the domain specificity perspective, abilities in each domain could be relatively independent of each other. As stated earlier, this could be an important reason why the OSPAN, which is more quantitatively loaded, was not as predictive of the more verbal-oriented task used in the current study. There is some support for WM as domain specific (e.g., Shah & Miyake, 1996; Shipstead & Yonehiro, 2016).

However, Kane et al. (2007), are proponents of the domain generality of WM, and they argue these findings are produced as a result of range restriction in cognitive ability when using university students and the use of only one measure of each specific type of WM task; both of

these conditions were existent in the current study, as university students were used and only one verbal WM task and one quantitative WM task were used.

Practical Implications

Past research has shown that assessment of specific cognitive abilities does not improve the predictive accuracy of “g” (Ree & Earles, 1991; Ree, Earles & Teachout, 1994; Stanhope & Surface, 2014). Thus, research on specific abilities had faded away and it is widely accepted that “g” alone is sufficient to use in the assessment battery given to applicants (Murphy, 2017; Schmidt & Hunter, 2004). Although more WM research is needed, the cumulative evidence that supports WM indicates that WM assessments create utilities that warrant consideration in personnel selection test batteries.

These results suggest that adding WM to currently used personnel selection measures of “g” could lead to increases in employee performance, which are important for organizations in both the public and private sector. Increments in predictive validity (i.e., 3%) has been suggested to have the possibility of saving the Navy \$83 million when added to the current test battery for selection (Wolfe et al., 1995).

Although the current measures of WM are more costly and time-consuming than measures of “g”, selection for some job types may warrant the use of WM measures. For example, air traffic controllers have a highly complex job that involves the safety of many people. In this case, the performance of air traffic controllers is of utmost importance, and the cost of poor performance is extremely high. It is selection for these types of jobs that WM measures are both justified and necessary. On the other hand, it does not seem as if organizations would use WM to assess applicants of lower-level jobs or jobs that are less complex. The use of

WM measures in this case would probably demand more money and resources than is necessary for the prediction of job performance.

Limitations and Future Directions

As briefly discussed earlier, one limitation of this study is the use of only one measure of WM for the two different types of WM (i.e., verbal and quantitative). This was not preemptively considered an issue because of the literature which supports WM as a domain general construct (Kane et al., 2007; Li, Christ, & Cowan, 2014). The length of this study was already 90 minutes, so adding more measures would have likely elicited fatigue among participants. However, if possible, future research should be conducted using more than one measure of different domains of WM, if domain generality is an issue in the model or if it is a research question being examined.

Future research can benefit from using other domains in the measurement of task performance. For example, adding a quantitatively loaded task in the battery may reveal that the OSPAN better predicts quantitatively loaded tasks, and RSPAN verbally loaded tasks (e.g., the current study). If WM is domain specific, for which this study provides some support, there are important implications for using domain specific WM measures in the organizational sciences, where personnel selection measures may change as a function of the job's responsibilities.

In most cases in the organizational sciences, field data is preferable over laboratory data. The time commitment necessary to participate (i.e., 90 minutes) and the computer programs used to administer the WM measures limited this study in that it would have been difficult and impractical to get organizations to participate in order to collect real world data.

Real world data would help to bridge the gap between task knowledge/task performance and job knowledge/job performance. This study, because it uses students and laboratory data, is

tasked with making a theoretical jump, proposing that task knowledge would translate to job knowledge and task performance would translate to job performance in organizational settings. However, the link between the two pairs of constructs is intuitive.

Another future direction for WM research should be to examine the subgroup differences associated with WM. Organizational practitioners are sometimes hesitant to use measures of “g” because they are found to result in subgroup differences, such that African-Americans score lower than Whites, and Whites score lower than Asian/Pacific Islanders (Gottfredson, 1988; Hunter & Hunter, 1984; Schmitt, Clause, & Pulakos, 1996). The sample sizes were not adequate to examine race differences in this study, but there is a small amount of research that has been done examining race differences. Thus far, WM seems to exhibit smaller racial/ethnic differences than Gf or general cognitive ability tests (Dehn, 2011; Raiford & Coalson, 2014; Weiner, Lerner, Easterbrooks, & Mistry, 2012), which is probably due to the fact that WM measures typically capture more processing functions which are less affected by race/ethnicity (Dehn, 2011). The absence of subgroup differences is important in many fields, and especially important to practitioners in the organizational sciences when examining diversity and adverse impact.

A key direction in future research should focus on parsing the effects of the capacity (WMC) and processing (WMP) functions of WM. Because they exist as a cognitive tradeoff (Daneman & Carpenter, 1980), it is difficult to administer a task that will allow researchers to examine the effects of one independent of the other. However, in order to more precisely test the theory underlying this partial mediation model, researchers must innovate new ways to measure WM in regards to capacity and processing because currently the most predominantly used

measures (i.e., OSPAN and RSPAN; Wilhelm, Hildebrandt, & Oberauer, 2013) do not measure these independently.

Lastly, if further research reveals WM as a domain-specific construct, selection for different job domains may require domain-specific WM abilities, such as spatial WM for air traffic controllers (Stein & Garland, 1993). The Army and Navy have already started to incorporate WM into both their research and personnel selection across various job domains (ASVAB, 2016), examining how different measures of WM may predict performance for different types and complexities of jobs.

Conclusions

The current study examined the specific cognitive ability of WM and its relationships to “g”, task knowledge, and task performance. This study provided a theory of underlying processes that contribute to the effect of WM on both task knowledge and task performance. The conflicting results of the OSPAN and RSPAN, as well as the predictive validity of the RSPAN in predicting task performance, open new avenues for future research. The current study provides initial evidence for the importance of considering cognitive abilities other than the “gold standard” for personnel selection batteries – “g”.

References

- Aben, B., Stapert, S., & Blokland, A. (2012). About the distinction between working memory and short-term memory. *Frontiers in Psychology*, 3:301.
- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, 131, 30 – 60.
- Ackerman, P. L., & Beier, M. E. (2007). Further exploration of perceptual speed abilities in the context of assessment methods, cognitive abilities, and individual differences during skill acquisition. *Journal of Experimental Psychology: Applied*, 13, 249-272.
- Alloway, T.P. (2009). Working memory, but not IQ, predicts subsequent learning in children with learning difficulties. *European Journal of Psychological Assessment*, 25, 92-98.
- Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, 106, 20–29.
- Armed Services Vocational Aptitude Battery (ASVAB) Subtests (2016). Retrieved from http://official-asvab.com/subtests_coun.htm.
- Babcock, R. L. (1994). Analysis of adult age differences on the Raven's advanced progressive matrices test. *Psychology and Aging*, 9, 303 – 314.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 47-89). San Diego, CA: Academic Press.
- Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*, 49, 5–28.
- Baddeley, A. D. (1999). *Human memory*. Boston: Allyn & Bacon.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, 4, 417–423.
- Baddeley, A. D., Eysenck, M., & Anderson, M. C. (2009). *Memory*. Hove: Psychology Press.

- Baddeley, A. (2010). Working memory. *Current Biology*, 20, 136–140.
- Baddeley, A. D., Allen, R. J., & Hitch, G. J. (2011). Binding in visual working memory: The role of the episodic buffer. *Neuropsychologia*, 49, 1393–1400.
- Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, 63, 1–29.
- Barrett, G. V., Polomsky, M. D., & McDaniel, M. A. (1999). Selection tests for firefighters: A comprehensive review and meta-analysis. *Journal of Business and Psychology*, 13, 507–513.
- Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences*, 29, 109-160.
- Bloom, L. C., & Watkins, M. J. (1999). Two-component theory of the suffix effect: Contrary findings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1452–1475.
- Bosco, F., Allen, D. G., & Singh, K. (2015). Executive attention: An alternative perspective on general mental ability, performance, and subgroup differences. *Personnel Psychology*, 68(4), 859–898.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY: Cambridge University Press.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston, MA: Houghton Mifflin.
- Cattell, R.B. (1987). *Intelligence: Its structure, growth, and action*. New York: Elsevier.
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A

- theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404–431.
- Chuderski, A. (2014). Relational integration task explains fluid reasoning above and beyond other working memory tasks. *Memory & Cognition*, 42, 448–463.
- Colom, R., Flores-Mendoza, C., & Rebollo, I. (2003). Working memory and intelligence. *Personality and Individual Differences*, 34, 33-39.
- Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user’s guide. *Psychonomic Bulletin & Review*, 12, 769 –786.
- Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. N. (Eds.) (2007). *Variations in working memory*. Oxford: Oxford University Press.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.
- Darwin, C. J., Turvey, M. T., & Crowder, R. G. (1972). An auditory analogue of the Sterling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology*, 3, 255-267.
- Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford, England: Oxford University Press.
- Dehn, M. J. (2008). *Working memory and academic learning: Assessment and intervention*. Hoboken, NJ: John Wiley & Sons, Inc.
- de Jong, P.F., & Das-Smaal, E.A. (1995). Attention and intelligence: The validity of the star counting test. *Journal of Educational Psychology*, 87, 80–92.
- DeKeyser, R. M., & Koeth, J. (2011). Cognitive aptitudes for second language learning. In E.

- Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 395-406). London: Routledge.
- Denis, M., Logie, R. H., Cornoldi, C., de Vega, M., & Engelkamp, J. (2001). *Imagery, language, and visuo-spatial thinking*. Hove, UK: Psychology Press.
- Driskell, J. E., & Olmstead, B. (1989). Psychology and the military: Research applications and trends. *American Psychologist*, *44*(1), 43-54.
- Duffau, H., Gatignol, P., Denvil, D., Lopes, M., Capelle, L. (2003). The articulatory loop: study of the subcortical connectivity by electrostimulation. *Neuroreport*, *14*, 2005-2008.
- Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in the role of working memory in comprehension and following directions. *Journal of Educational Research*, *84*, 253-262.
- Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). 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. (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp.102-134). London: Cambridge Press.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309–31.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19-23.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The Psychology of Learning and*

- Motivation* (Vol. 44, pp. 145-199). NY: Elsevier.
- Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211-245.
- Filičková, M., Ropovik, I., Bobaková, M., & Kovalčíková, I. (2015). The relationship between fluid intelligence and learning potential: Is there an interaction with attentional control? *Journal of Pedagogy*, *6*(1), 25-41.
- Flanagan D. P., & Dixon S. G. (2013). The Cattell–Horn–Carroll theory of cognitive abilities. In Reynolds, C. R., Vannest, K. J., & Fletcher-Janzen, E. (Eds.), *Encyclopedia of special education* (pp. 368–382). Hoboken, NJ: John Wiley & Sons.
- Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition*, *43*, 226–236.
- Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*, *54*, 1-34.
- Ganzach Y. (1998). Intelligence and job satisfaction. *Academy of Management Journal*, *41*, 526–539.
- Gathercole, S. E., & Alloway, T. P. (2004). Working memory and classroom learning. *Dyslexia Review*, *15*, 4-9.
- Gathercole, S. E., Lamont, E., & Alloway, T. P. (2006). Working memory in the classroom. In Pickering, S. (Ed.), *Working memory and education* (pp. 219–240). Oxford, UK: Elsevier.
- Gersten, R., Jordan, N.C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*, *38*, 293–304.

- Ghiselli, E., & Brown, C. (1951). The Placement of workers: Concepts and problems. *Personnel Psychology*, 9(1), 1-16.
- Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology*, 26(4), 461-477.
- Gignac, G. E. (2014). Fluid intelligence shares closer to 60% of its variance with working memory capacity and is a better indicator of general intelligence. *Intelligence*, 47, 122–133.
- Gottfredson, L. S. (1988). Reconsidering fairness: A matter of social and ethical priorities. *Journal of Vocational Behavior*, 33, 293-319.
- Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24, 79 –132.
- Guttman, N., & Julesz, B. (1963). Lower limits of auditory periodicity analysis. *Journal of the Acoustical Society of America*, 35, 610.
- Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology*, 24, 1149–1167.
- Held, J. D., & Wolfe, J. H. (1997). Validities of unit-weighted composites of the Armed Services Vocational Aptitude Battery (ASVAB) and Enhanced Computer Administered Tests (ECAT). *Military Psychology*, 9, 77-84.
- Hirsh, H. R., Northrop, L. C., & Schmidt, F. L. (1986). Validity generalization results for law enforcement occupations. *Personnel Psychology*, 39, 399-420.
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, 57, 253-270.

- Horn, J. L. (1991). Measurement of intellectual capabilities: A review of theory. In McGrew, K. S., Werder, J.K., & Woodcock, R. W. (Orgs.), *Woodcock-Johnson technical manual* (pp. 197- 232). Chicago: Riverside
- Horn, J. L. (1998). A basis for research on age differences in cognitive abilities. In McArdle, J. J., & Woodcock, R. W. (Eds.), *Human cognitive abilities in theory and practice* (pp. 57 – 92). Mahwah, NJ: Erlbaum.
- Hummel, K. M. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics*, 30, 225–249.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- Hunter, J. E. (1986). Cognitive ability, cognitive aptitude, job knowledge, and job performance. *Journal of Vocational Behavior*, 29(3), 340-362.
- Jensen, A. R. (1986). G: Artifact or reality? *Journal of Vocational Behavior*, 29, 301-331.
- Jensen, A. R. (1998). *The g factor: the science of mental ability*. Westport, CT: Praeger.
- Jensen, A. R. (2002). Psychometric g: Definition and substantiation. In R. J. Sternberg & E. L. Grigorenko (Eds.), *General factor of intelligence: How general is it?* (pp. 39 –54). Mahwah, NJ: Erlbaum.
- Judge, T. A., Colbert, J. E., & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology*, 89, 542-552.
- Judge, T. A., Klinger, R. L., & Simon, L. S. (2010). Time is on my side: time, general mental ability, human capital, and extrinsic career success. *Journal of Applied Psychology*, 95, 92–107.
- Jurden, F. H. (1995). Individual differences in working memory and complex cognition. *Journal*

of Educational Psychology, 87, 93–102

- Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, 9, 637-671.
- Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory*. New York, NY: Oxford.
- Kaufman, S. B. (2009). *Beyond general intelligence: The dual-process theory of human intelligence* (Unpublished Ph.D. dissertation). Yale University, New Haven, CT.
- Kieras, D. E., Meyer, D. E., Mueller, S. T., & Seymour, T. L. (1999). Insights into working memory from the perspective of the EPIC architecture for modeling skilled perceptual-motor and cognitive human performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 183-223). New York, NY, US: Cambridge University Press.
- Kiewra, K. A., & Benton, S. L. (1988). The relationship between information-processing ability and notetaking. *Contemporary Educational Psychology*, 13, 33-44.
- King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30, 580-602.
- Krumm, S., Ziegler, M., & Buehner, M. (2008). Reasoning and working memory as predictors of school grades. *Learning and Individual Differences*, 18(2), 248–257.
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of*

Personality and Social Psychology, 86, 148–161

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success.

Science, 315(5815), 1080-1081.

Kyllonen, P., & Christal, R. (1990). Reasoning ability is (little more than) working memory

capacity?! *Intelligence*, 14, 389–433.

Kyllonen, P., & Stephens, D. L. (1990). Cognitive abilities as determinants of success

in acquiring logic skill. *Learning and Individual Differences*, 2, 129-160.

Li, D., Christ, S. E., & Cowan, N. (2014). Domain-general and domain-specific functional

networks in working memory. *NeuroImage*, 102, 646–656.

Linck, J., & Weiss, D.J. (2015) Working memory predicts second language classroom learning.

Sage Open, 5(4), doi: 10.1177/2158244015607352.

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994) Counting on working memory in arithmetic

problem solving. *Memory and Cognition*, 22, 395–410.

Lord, R. G., deVader, C. L., & Alliger, G. M. (1986). A meta-analysis of the relation between

personality traits and leadership perceptions: An application of validity generalization procedures. *Journal of Applied Psychology*, 71, 402-410.

Luo, D., Thompson, L.A., & Detterman, D. K. (2006). The criterion validity of tasks of basic

cognitive processes. *Intelligence*, 34, 79–120.

Matthews, T., & Lassiter, K. (2007). What does the Wonderlic Personnel Test measure?

Psychological Reports, 100, 707-712.

McGrew, K. S. (2005). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan

& P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 136–181). New York, NY: Guilford Press.

- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, 1-10.
- Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review, 63*, 81-97.
- Miller, G. A., Galanter, E., Pribram, K. H. (1960). *Plans and the structure of behavior*. New York, NY: Henry Holt & Company.
- Miyake, A., & Shah, P. (1999). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 442–481). New York, NY: Cambridge University Press.
- Mount, M. K., Oh, I. S., & Burns, M. (2008). Incremental validity of perceptual speed and accuracy over general mental ability. *Personnel Psychology, 61*(1), 113-139.
- Murphy, K. (2017). What can we learn from “not much more than g”? *Journal of Intelligence, 5*(8), 1-6.
- Oberauer, K., Süß, H. M., Wilhelm, O., & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence, 31*, 167-193.
- Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H. M. (2005). Working memory and intelligence—their correlation and their relation: A comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*, 61–65.
- Olea, M., & Ree, M. J. (1994). Predicting pilot and navigator criteria: Not much more than g. *Journal of Applied Psychology, 79*, 845-849.
- Ones, D. S., Dilchert, S., & Viswesvaran, C. (2012). Cognitive abilities. In N Schmitt (Ed.) *The*

- Oxford handbook of personnel assessment and selection*, (pp.179-224). New York: Oxford University Press.
- Palumbo, M. V., Miller, C. E., Shalin, V. L., & Steele-Johnson, D. (2005). The impact of job knowledge in the cognitive ability-performance relationship. *Applied H.R.M. Research*, *10*(1), 13-20.
- Pawlak, W.S., Brinton, C.S., Crouch, K., & Lancaster, K.M. (1996). A framework for the evaluation of air traffic control complexity. Presented at the AIAA national Conference.
- Pearlman, K., Schmidt, F. L., & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. *Journal of Applied Psychology*, *65*, 373-406.
- Raiford, S. E., & Coalson, D. L. (2014). *Essentials of WPPSI-IV assessment*. Hoboken, NJ: John Wiley & Sons, Inc.
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General overview*. Oxford, England: Oxford Psychologists Press.
- Ree, M. J., Earles, J. A. (1991). Predicting training success: Not much more than g. *Personnel Psychology*, *44*, 321-332.
- Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science*, *1*, 86-89.
- Ree, M. J., Earles, J. A., Teachout, M. S. (1994). Predicting job performance: Not much more than g. *Journal of Applied Psychology*, *79*, 518-524.
- Ren, X., Schweizer, K., Wang, T., & Xu, F. (2015). The prediction of students' academic performance with fluid intelligence in giving special consideration to the contribution of

- learning. *Advances in Cognitive Psychology*, 11, 97–105.
- Repov, G., & Baddeley, A. D. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience*, 139, 5–21.
- Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High-stakes testing in higher education and employment: Appraising the evidence for validity and fairness. *American Psychologist*, 63, 215–227.
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F., & Rolland, J. P. (2003). A meta-analytic study of general mental ability validity for different occupations in the European community. *Journal of Applied Psychology*, 88, 1068-1081.
- Salthouse, T. A., Mitchell, D. R. D, Skovronek, E., & Babcock, R. L. (1989). Effects of adult age and working memory on reasoning and spatial abilities. *Journal of Experimental Psychology Learning: Memory & Cognition*, 15, 507-516.
- Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, 2, 179-183.
- Salthouse, T. A., & Pink, J. E. (2008). Why is working memory related to fluid intelligence? *Psychonomic Bulletin & Review*, 15, 364–371.
- Sams, M., Hari, R., Rif, J., & Knuutila, J. (1993). The human auditory sensory memory trace persists about 10s: Neuromagnetic evidence. *Journal of Cognitive Neuroscience*, 5, 363–70.
- Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, 71(3), 432-439.
- Schmidt, F. L., Hunter, J. E., & Dunn, W. L. (1987). Potential utility increases from adding new

- tests to the Armed Services Vocational Aptitude Battery (ASVAB). *Navy Personnel Research and Development Center*, San Diego, CA: Unpublished manuscript.
- Schmidt, F. L., & Hunter, J. E. (1992). Development of causal models of processes determining job performance. *Current Directions in Psychological Science*, *1*, 89-92.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262-274.
- Schmidt, F. L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, *15*(1-2), 187-211.
- Schmidt, F. L., & Hunter J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162-173
- Schmitt, N., Clause, C., & Pulakos, E. (1996). Subgroup differences associated with different measures of some common job relevant constructs. In Cooper CL, Robertson IT (Eds.), *International review of industrial and organizational psychology*, *11*, 5-137.
- Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues*, 3rd ed., pp. 99-144). New York, NY: Guilford Press.
- Shipstead, Z., & Yonehiro, J. (2016). The domain-specific and domain-general relationships of visuospatial working memory to reasoning ability. *Psychonomic Bulletin and Review*, 1-9.
- Spearman, C. (1904). General intelligence, objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201-293.

- Spearman, C. (1946). Theory of general factor. *British Journal of Psychology*, 36, 117–131.
- Stanhope, D., & Surface, E. (2014). Examining the incremental validity and relative importance of specific cognitive abilities in a training context. *Journal of Personnel Psychology*, 13(3), 146-156.
- Stauffer, J. M., Ree, M. J., & Caretta, T. R. (1996). Cognitive-components tests are not much more than g: An extension of Kyllonen's analysis. *Journal of General Psychology*, 123, 193–205.
- Stein, E., & Garland, D. (1993). Air traffic controller working memory: Considerations in air traffic control tactical operations. *National Technical Information Service Report*, 1-66.
- Swanson, H.L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. *Journal of Experimental Child Psychology*, 79, 294–321.
- Swanson, H.L., & Kim, K. (2007). Working memory, short-term memory, and naming speed as predictors of children's mathematical performance. *Intelligence*, 35, 151-168
- Thomas, S. L., & Scroggins, W. A. (2006). Psychological testing in personnel selection: Contemporary issues in cognitive ability and personality testing. *Journal of Business Inquiry*, 5, 28-38.
- Tucker, P., & Warr, P. (1996). Intelligence, elementary cognitive components, and cognitive styles as predictors of complex task performance. *Personality and Individual Differences*, 21, 91–102.
- Undheim, J. O., & Gustafsson, J. E. (1987). The hierarchical organization of cognitive

- abilities: Restoring general intelligence through the use of linear structural relations. *Multivariate Behavioral Research*, 22, 149- 171.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37, 498 –505.
- Vallar, G. (2006). Memory systems: The case of phonological short-term memory. A festschrift for cognitive neuropsychology. *Cognitive Neuropsychology*, 23, 135-155.
- Vernon, P. E., & Parry, J. B. (1949). *Personnel selection in the British Forces*. London: University of London Press.
- Viswesvaran, C., & Ones, D. S. (2002). Examining the construct of organizational justice: A meta-analytic evaluation of relations with work attitudes and behaviors. *Journal of Business Ethics*, 38, 193–203.
- Weiner, I. B., Lerner, R. M., Easterbrooks, M. A., & Mistry, J. (2012). *Handbook of psychology, developmental psychology*. Hoboken, NJ: John Wiley & Sons, Inc.
- Wen, Z., & Skehan, P. (2011). A new perspective on foreign language aptitude research: Building and supporting a case for “working memory as language aptitude.” *Revista Ilha do Desterro*, 60, 15–43.
- Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, 4. doi: 10.3389/fpsyg.2013.00433.
- Wilson, K. M., & Swanson, H. L. (2001). Are mathematics disabilities due to a domain-general or a domain-specific working memory deficit? *Journal of Learning Disabilities*, 34(3), 237–248.
- Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin & Review*, 8, 44-57.

Wolfe, J. H., Alderton, D. L., Larson, G. E., & Held, J. D. (1995). Incremental validity of enhanced computer-administered testing (ECAT). San Diego, CA: Navy Personnel Research and Development Center.

Yuan, K., Steedle, J., Shavelson, R., Alonzo, A., & Oppezzo, M. (2006). Working memory, fluid intelligence, and science learning. *Educational Research Review, 1*, 83–98.

Appendices

Appendix A – Journal Article Information for Task Performance Measure

Please use the following information about five **journal articles** and create a properly formatted APA style reference page. Based on the tutorial you just watched, try to be as accurate as possible including proper *formatting, punctuation, capitalization, spelling, italics, grammar, and spacing*. The APA reference page that has the most accurate formatting and the least mistakes will win a \$50 Gift Card to Belk®. You will have 12 minutes to complete this task.

1. Journal: European Journal of Marketing

Pages of article in journal: Pages 1245-1283

Volume of journal: Volume 41

Published: November 17, 2007

Author(s): Douglas Brownlie

Title of article: Toward effective poster presentations: An annotated bibliography

2. Journal: Group Processes & Intergroup Relations

Pages of article in journal: Pages 28-45

Volume of journal: Volume 16

Published: January 10, 2013

Author(s): Emery Cecile, Thomas S. Calvard, Meghan E. Pierce

Title of article: Leadership as an emergent group process: A social network study of personality and leadership

3. Journal: Psychological Review

Pages of article in journal: Pages 369-389

Volume of journal: Volume 94

Published: June 16, 1987

Author(s): Kay Deaux, Brenda Major

Title of article: Putting gender into context: An interactive model of gender-related behavior

4. Journal: Academy of Management Review

Pages of article in journal: Pages 889-913

Volume of journal: Volume 31

Published: September 24, 2006

Author(s): Piers Morgan, Cornelius J. Konig

Title of article: Integrating theories of motivation

5. Journal: Journal of Personality and Social Psychology

Pages of article in journal: Pages 1173-1182

Volume of journal: Volume 51

Published: March 26, 1986

Author(s): Reuben M. Baron, David A. Kenny

Title of article: The Moderator-Mediator Variable Distinction in Social Psychological Research:
Conceptual, Strategic, and Statistical Considerations

Appendix B – APA Pretest

1. How are entries ordered in a Reference list?

- A. Title, author's name, date of publication, journal, volume, pages.
- B. Author's name, date of publication, title, journal, volume, pages.**
- C. Date of publication, author's name, title, journal, pages, volume.
- D. Author's name, title, date of publication, journal, pages, volume.

2. Which of the following is the correct APA Reference list entry?

- A. Reese, G. (2000). *Database Programming with JDBC and Java*. Beijing: O'Reilly Media.**
- B. Reese, George. *Database programming with JDBC and Java*. Beijing: O'Reilly Media. 2000.
- C. Reese, G. (2000). *Database programming with JDBC and Java*. O'Reilly Media, Beijing, China.
- D. Reese, G. (July 2000). *Database programming with JDBC and Java*. O'Reilly Media, Beijing, China.

3. Below is a reference page listing for a book with two authors. How should the title of the book be formatted?

Giorgis, C., & Glazer, J. I. (2009). Literature for young children: Supporting emergent literacy, ages 0 - 8 (6th ed). Boston, MA: Pearson Education.

- A. The book title should be in quotations.
- B. The book title should be in italics.**
- C. The book title should be listed without special formatting.
- D. The book title should be underlined.

4. Select the answer that provides the correct reference page format for a journal article published online and in print.

- A. Recent work on epistemic value. *American Philosophical Quarterly*, 44 (2), 85-110.
Retrieved from <http://www.jstor.org/stable/20464361>
- B. Pritchard (2007). Recent work on epistemic value. Retrieved from
<http://www.jstor.org/stable/20464361>
- C. **Pritchard, D. (2007, April). Recent work on epistemic value. *American Philosophical Quarterly*, 44 (2), 85-110. Retrieved from <http://www.jstor.org/stable/20464361>**
- D. Pritchard, D. (2007, April). *American Philosophical Quarterly*, 44 (2), 85-110. Retrieved from <http://www.jstor.org/stable/20464361>

5. In the following example of a reference page listing of a journal article retrieved from an electronic database, which item(s) of information are missing?

Florian. (2010). Challenges for interactivist-constructivist robotics. *New Ideas in Psychology*. 350-353. doi: 10.1016/j.newideapsych.2009.09.009

- A. The DOI, the month, and the page number
- B. The author's initials and the page number
- C. **The author's initials, the journal volume and issue number**
- D. The month, volume, and issue number

6. In what order should your references be listed on the reference page?

- A. Chronologically from oldest to most recent
- B. Alphabetically by title of article
- C. The order in which they appear in your paper
- D. **Alphabetically by author**

7. Which one of the following references is properly formatted?

- A. Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology*, 35, 127-145.
- B. Alibali, M., W. (1999). How Children Change Their minds: Strategy Change Can Be Gradual or Abrupt. *Developmental Psychology*, 35, 127-145.
- C. Alibali, M. W. (1999). *How children change their minds: Strategy change can be gradual or abrupt*. *Developmental Psychology*, 35, pgs. 127-145.
- D. **Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology*, 35, pgs. 127-145.**

8. What is wrong with the following APA style journal reference?

Swan, J. T., Rail, D. Q., & Bushcombe, M. N. (1995). "Students and the problem of entitlement." *Social Work and Research*, 68(2), 127-137.

- A. **The article's title should not be in quotations.**
- B. The journal title should not be in parentheses.
- C. The important words in the article title should be capitalized.
- D. The date should include the month the article was published.

9. What type of citation is this?

Austin, J. H. (1998). *Zen and the brain: Toward an understanding of meditation and consciousness*. Cambridge, MA: MIT Press.

- A. Journal article
- B. Magazine article
- C. Book chapter

- D. **Book**
- E. Newspaper article

10. An APA reference page should be formatted as follows:

- A. Single-spaced, 12-pt font, hanging indents, references in chronological order
- B. Double-spaced, 12-pt font, left-line indents, references in chronological order
- C. **Double-spaced, 12-pt font, hanging indents, references in order by author's last name**
- D. Single-spaced, 12-pt font, left-line indents, references in order by author's last name

APA Posttest:

1. When typing up a journal reference, how should the journal title be formatted?

- A. Written in italics, only important words capitalized, followed by a period.
- B. Written in bold, only first word capitalized, followed by a comma.
- C. No special formatting, only first word capitalized, followed by a period.

D. Written in italics, only important words capitalized, followed by a comma.

E. Written in bold, only first word capitalized, followed by a comma.

2. What is wrong about the following APA reference?

Jacoby, W. G. (1994). Public Attitudes Toward Government Spending. *American Journal of Political Science*, 38(2), 336-361.

- A. The title of the article should come before the author.
- B. Only the first word of the article title should be capitalized.**
- C. Only the first word of the journal title should be capitalized.
- D. The title of the article should be italicized, not the journal.
- E. The (2) in parentheses should not be used unless it is a book.

3. Which of the following is a properly formatted reference for the reference page?

A. Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War: an Historical Review. *American Political Science Review*, 97(01), 75-98. doi: 10.1017/S0003055403000534.

B. Fearon, J. D., & Laitin, D. D. (June 2003). Ethnicity, insurgency, and civil war: an historical review. *American Political Science Review*, 97(01), pgs. 75-98. doi: 10.1017/S0003055403000534.

C. Fearon, J. D., & Laitin, D. D. (June 2003). *Ethnicity, insurgency, and civil war: An historical review*. *American Political Science Review*, 97(01), 75-98. doi: 10.1017/S0003055403000534.

D. Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War: An Historical Review. *American Political Science Review*, 97(01), p. 75-98. doi: 10.1017/S0003055403000534.

E. Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, insurgency, and civil war: An historical review. *American Political Science Review*, 97(01), 75-98. doi: 10.1017/S0003055403000534.

4. What is the correct order of items to be included in a reference?

A. Author, article title, year published, journal, volume, pages.

B. Author, year published, article title, journal, volume, pages.

C. Author, journal, year published, article title, page numbers

D. Article title, author, year published, page numbers, journal.

E. Author, year published, journal, article title, pages.

5. In what order should your references be listed on the reference page?

A. Chronologically from oldest to most recent

B. Alphabetically by title of article

C. The order in which they appear in your paper

D. Alphabetically by author

6. What is wrong with the following APA style journal reference?

Tidwell, L. C., & Walther, J. B. (2002). Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: getting to know one another a bit at a time. *Human Communication Research*, 28, 317-348.

A. The word “getting” in the title should be capitalized.

B. The journal title should not be in italics, the article title should be.

C. The important words in the article title should be capitalized.

D. The date should include the month the article was published.

E. The initials and names of the authors are not formatted correctly.

7. How should the names of the authors be formatted in an APA reference?

A. B. Harrison, J. F. Wood, V. Smith, & N. Westbrook (2004).

B. Harrison B., Wood J. F., Smith V., and Westbrook N. (2004).

C. Harrison, B., Wood, J. F., Smith, V., & Westbrook, N. (2004).

D. Harrison, B., Wood, J. F., Smith, V., and Westbrook, N. (2004).

E. Harrison, Brenda; Wood, Jason Flores; Smith, Veronica; & Westbrook, Nick (2004).

8. When should periods appear in an APA reference page?

A. After authors' initials, after article title, and after page numbers.

B. After authors' initials, after year published, after article title, and after page numbers.

C. After authors' initials, after year published, after article title, after journal title, and after page numbers.

D. After authors' initials, after year published, and after page numbers.

E. After authors' initials, after article title, after volume numbers, and after page numbers.

9. When should commas appear in an APA reference page?

A. Between authors, after journal title, and after volume number.

B. After article title, after journal title, and after volume number.

C. Between authors, after article title, after journal title.

D. Between authors, and after journal title.

E. After journal title, after volume number, and after year published.

10. Which of the below should you NOT see in an APA journal reference?

- A. An italicized journal title with important words capitalized.
- B. An article title with only one capitalized word.
- C. The issue number of the journal's volume.
- D. The word "pgs." to explain you are listing the article's page numbers.**
- E. A capitalized word after a semicolon.

Appendix C

OSPAN. Example of one item from the OSPAN. Letters to be remembered were presented in blocks of three to seven.

The screenshot displays the OSPAN task interface, which is divided into four main sections:

- Top Left:** A grey box containing the math problem $(7/7) + 6 =$ and the instruction "Click the mouse to continue".
- Top Right:** A grey box containing the number "7" and two buttons labeled "TRUE" and "FALSE".
- Bottom Left:** A grey box containing the letter "R".
- Bottom Right:** A white box containing the instruction "Select the letters in order. Use the blank button to fill in forgotten letters". Below this instruction is a grid of letter selection options:

<input type="checkbox"/> F	<input type="checkbox"/> H	<input type="checkbox"/> J
<input type="checkbox"/> K	<input type="checkbox"/> L	<input type="checkbox"/> N
<input type="checkbox"/> P	<input type="checkbox"/> Q	<input type="checkbox"/> R
<input type="checkbox"/> S	<input type="checkbox"/> T	<input type="checkbox"/> Y

 Below the grid is a "blank" button. At the bottom of this section are "clear" and "Exit" buttons.

Appendix D – Scoring Rubric for APA Reference Page

Reference #	Criterion	Points scored
1	Participant spells everything correctly throughout the reference	
1	Participant includes periods in all the correct places.	
1	Participant places commas in the correct places throughout the reference	
1	Participant places parentheses in the correct places throughout the reference	
1	Participant properly capitalizes as needed throughout the reference	
1	Participant puts spaces in the correct places throughout the reference	
1	Participant puts necessary information in italics	
1	Participant places each item of information in the proper order in the entire reference	
2	Participant spells everything correctly throughout the reference	
2	Participant includes periods in all the correct places.	
2	Participant places commas in the correct places throughout the reference	
2	Participant places parentheses in the correct places throughout the reference	
2	Participant properly capitalizes as needed throughout the reference	
2	Participant puts spaces in the correct places throughout the reference	
2	Participant puts necessary information in italics	

2	Participant places each item of information in the proper order in the entire reference	
3	Participant spells everything correctly throughout the reference	
3	Participant includes periods in all the correct places.	
3	Participant places commas in the correct places throughout the reference	
3	Participant places parentheses in the correct places throughout the reference	
3	Participant properly capitalizes as needed throughout the reference	
3	Participant puts spaces in the correct places throughout the reference	
3	Participant puts necessary information in italics	
3	Participant places each item of information in the proper order in the entire reference	
4	Participant spells everything correctly throughout the reference	
4	Participant includes periods in all the correct places.	
4	Participant places commas in the correct places throughout the reference	
4	Participant places parentheses in the correct places throughout the reference	
4	Participant properly capitalizes as needed throughout the reference	
4	Participant puts spaces in the correct places throughout the reference	
4	Participant puts necessary information in italics	

4	Participant places each item of information in the proper order in the entire reference	
5	Participant spells everything correctly throughout the reference	
5	Participant includes periods in all the correct places.	
5	Participant places commas in the correct places throughout the reference	
5	Participant places parentheses in the correct places throughout the reference	
5	Participant properly capitalizes as needed throughout the reference	
5	Participant puts spaces in the correct places throughout the reference	
5	Participant puts necessary information in italics	
5	Participant places each item of information in the proper order in the entire reference	
	Participant puts all references in the correct order on the reference page	
Total Score:	Sum of the points in the boxes to the right.	

Appendix E – Script for the lab study

“This is a cognition and learning study. Therefore, there will be two parts to this experiment: one will involve cognitive tasks, and the other will involve learning tasks. We will be starting each task together, so if you finish any given task before others, just sit tight and relax until we start the next task. Once we start, you will complete a series of cognitive tasks with breaks in between them. Please do your best on these tasks, and try to have some fun with them. The best performer across all tasks will receive a \$50 cash prize, and the 2nd best will receive a \$20 cash prize. All scores will be emailed to you at the conclusion of the study. Before we get started, does anybody have any questions for me? Now, we are ready to begin the first task, which will last about 25 minutes. You can all click the “START” button on your screen and input your participant number and click the next arrow. You can follow the instructions from here on out until you are done with this first task. Again, if you finish early, just sit tight for the start of the next task.”

~25 mins will go by...

“You will now complete a puzzle task that gives you a set of pictures, and your job is to determine which pattern or shape comes next in the sequence of pictures. Choose the option that best represents your answer. You will have 20 minutes to complete the 20 questions. Once you are done, sit tight until everyone has finished.”

~20 mins...

“You will now start the last cognitive portion of this study. It will be very similar to the first task you completed. It will last about 25 minutes just like the first task. When you are ready, please click the “START” button to get started.”

~25 mins....

“We will now take a break. For this break, I want you to tilt your head back, close your eyes, relax, and think of a calm, relaxing environment. You will have 5 minutes, but I want you to relax during this 5 minutes.”

“Okay, you are now done with the cognitive portion of this study and will be moving onto the learning part. As the instructions on your page read, you will have 6 minutes to answer 10 quiz questions related to APA reference page formatting. Do the best you can on these. I will be timing your 6 minutes this time, and will keep you updated with the time left. Once you click the next arrows, your time will start. You can go ahead and click the next arrows to begin.”

“4 minutes remaining”

“2 minutes remaining”

“You will now watch a 10 minute tutorial about APA formatting. You will learn formatting rules as well as how to create your own APA reference page. Please put the headphones on so you are able to hear the video. Once the video is over, we will be taking a short knowledge test to see how much you have learned. You may now start the video.”

“You guys have been doing great so far. You only have two more tasks left in the experiment. The first task will involve a short 10-question quiz on the information you have just learned. As the instructions

say, you will have 6 minutes to complete a 10-question multiple choice quiz once you click the next arrows. Please click the next arrows, put in your Participant ID, and begin the quiz. I will keep you updated on time remaining.”

“4 minutes remaining”

“2 minutes remaining”

“You are now ready to begin the final task. I will be giving you article information for 5 journal articles. Your job is to create a properly formatted APA reference page based on the information you have learned in the video and from the knowledge tests. Please pay careful attention to the formatting rules and guidelines. You will have 12 minutes to put all 5 articles into a reference page. Once I hand you the information, you may begin.”

“8 minutes remaining”

“4 minutes remaining”

“1 minute remaining”

“Okay, that’s all of the tasks we have for you today. Please either click the save icon on the top left or type Control + S. You can leave the Word document up and leave the papers where you are, and gather your belongings. We will be sending out emails after the study giving you feedback on your scores as well as notifying you if you have won either of the prizes. Thank you and have a great day.”