A Meta-Analysis of Information Processing Measures of Intelligence, Performance, and Group Score Differences

Elliott C. Larson
The Graduate Center, City University of New York

How does access to this work benefit you? Let us know!
Follow this and additional works at: https://academicworks.cuny.edu/gc_etds
Part of the Industrial and Organizational Psychology Commons

Recommended Citation
https://academicworks.cuny.edu/gc_etds/3040

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
A META-ANALYSIS OF INFORMATION PROCESSING MEASURES OF INTELLIGENCE, PERFORMANCE, AND GROUP SCORE DIFFERENCES

by

ELLIOTT CROFTS LARSON

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements of the degree of Doctor of Philosophy, The City University of New York

2019
A META-ANALYSIS OF INFORMATION PROCESSING MEASURES OF INTELLIGENCE, PERFORMANCE, AND GROUP SCORE DIFFERENCES

by

Elliott Crofts Larson

The manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirements for the degree of Doctor of Philosophy

Charles Scherbaum

__________________                                    ________________________________________

Date                                    Chair of Examining Committee

__________________

Richard Bodnar

__________________

Date                                    Executive Officer

__________________

Harold Goldstein

Logan Watts

Kristin Sommer

Kenneth Yusko

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK
ABSTRACT

A Meta-Analysis of Information Processing Measures of Intelligence, Performance, and Group Score Differences

by

Elliott Crofts Larson

Advisor: Dr. Charles A. Scherbaum

Intelligence is one of the most studied constructs in industrial-organizational (I-O) and educational psychology. Findings from numerous studies and meta-analyses have consistently demonstrated the power of intelligence measures to predict performance across a wide range of domains. This research has been fruitful and provides strong evidence for the utility of intelligence measures in organizations and schools. However, while intelligence measures have been developed and applied for over a century, most research in I-O psychology has relied on operationalizations of intelligence that focus on a person’s knowledge. Meta-theories of intelligence propose that intelligence can simultaneously be conceptualized as a person’s ability to process information. From this perspective, intelligence is not just what a person knows but also a person’s ability to maintain, learn, and use information to reason. Approaching intelligence as information processing offers unique opportunities for assessing intelligence that may explain additional variance in workplace outcomes in conjunction with commonly used intelligence measures. Furthermore, theory and early data suggests information processing measures may reduce group score differences typically reported with other types of intelligence measures. For these reasons, it is important to understand the existing literature on information processing measures. The current study offers insight into the utility of information processing measures in applied settings through a meta-analytic design. Samples examining information processing measures were collected and examined to evaluate if they predict job performance,
job-oriented training performance, and academic performance. Several variables, including the theoretical approach used to develop the measure, the diversity of task types in the measure, and the language knowledge requirement needed to respond to items on the measure, were tested as moderators of these relationships. In addition, the group score differences between African Americans and Caucasians on information processing measures were analyzed. Overall, the findings support information processing measures as valid predictors of outcomes across several critical domains. Group score differences were also found to be smaller than estimates from prior meta-analyses examining other intelligence measures. Results from the moderator tests provided several interesting trends in the data that require further examination in future research. The current study offers insights into how information processing measures operate in applied settings and is a step towards expanding the use of these measures in I-O psychology.
Acknowledgements

This project was only possible with the continuous and unconditional support from my family, friends, and colleagues. I am forever indebted to all of you. Foremost, I would not be here if it was not for my advisor, Charles, whose guidance, patience, and generosity were a constant source of encouragement. I would also like to thank my committee, Harold, Logan, Kristin, and Ken, for their time and effort in refining my dissertation with their questions, suggestions, and expertise. To my cohort and fellow graduate students, thank you for the friendships that made the journey through grad school far more enjoyable and manageable. To my loving family, I am grateful for everything you sacrificed to get me here. And most of all, to Clarice, thank you for your daily affirmations. You are my greatest inspiration and have motivated me to do things that at times seem impossible.
Table of Contents

Introduction ........................................................................................................................................ 1
Literature Review ................................................................................................................................. 6
  Intelligence ....................................................................................................................................... 6
  Measuring Intelligence-as-Knowledge ............................................................................................... 15
  Measuring Intelligence-as-Process ...................................................................................................... 20
  Hypotheses ....................................................................................................................................... 46
Method ................................................................................................................................................. 60
  Operationalization of Key Variables ................................................................................................. 61
Literature Search .................................................................................................................................. 66
  Inclusion and Exclusion Criteria ......................................................................................................... 67
  Coding ........................................................................................................................................... 69
Discussion ............................................................................................................................................ 93
  Overview of Findings ......................................................................................................................... 94
  Theoretical and Practical Implications ............................................................................................... 103
  Limitations ....................................................................................................................................... 109
  Future Research ................................................................................................................................. 110
  Conclusion ....................................................................................................................................... 113
Appendix A: Coding Form .................................................................................................................... 114
References .......................................................................................................................................... 139
List of Tables and Figures

Table 1. Representative Sample of Psychometric Information Processing Measures .............. 119
Table 2. Representative Sample of Cognitive Science Information Processing Measures ....... 121
Table 3. Reliability Coefficients Used for Correcting Measurement Error in the Criterion ...... 122
Table 4. Information Processing Measures and Performance Outcomes ................................ 123
Table 5. Group Score Differences ........................................................................................ 124
Table 6. Theoretical Approach and Performance Outcomes ................................................. 125
Table 7. Group Score Differences by Theoretical Approach .................................................. 126
Table 8. Task Diversity Within the Psychometric Approach and Performance Outcomes ...... 127
Table 9. Task Diversity Within the Cognitive Sciences Approach and Performance Outcomes .............................................................................................................. 128
Table 10. Language Knowledge Requirements and Performance Outcomes ......................... 129
Table 11. Summary of Findings by Hypothesis ..................................................................... 131
Figure 1. Decision tree for inclusion/exclusion criteria of meta-analysis ......................... 138
Intelligence has always been considered a complex construct with a wide range of definitions (Sternberg & Detterman, 1987). These definitions vary in the attributes they emphasize, such as those that frame intelligence around high-level cognitive functions whereas others are more honed on intelligence as knowledge structures (e.g., Jensen, 1980; Neisser et al., 1996; Sternberg, 2003; Sternberg & Detterman, 1987). To better map the domain of intelligence, Ackerman (1996) offered a meta-theory of intelligence that distinguishes between intelligence-as-process and intelligence-as-knowledge. This model effectively recognizes that intelligence is not a singular construct, but rather has various factors to it. While the idea that there are multiple domains of intelligence is not a new one (e.g., Cattell-Horn-Carroll model) (McGrew, 1997), Ackerman’s meta-theory is particularly effective at demonstrating the need for different types of measures. Specifically, this meta-theory emphasizes that certain measures of intelligence are better suited to assess people’s ability to process information or to assess their existing knowledge structures, and the type of measure applied in a situation should be aligned with the testing needs.

Ackerman’s meta-theory is a useful tool for summarizing current conceptualizations of intelligence, but also highlights the gaps that still exist in how intelligence is examined (Ackerman & Beier, 2005). While advancements in several disciplines, including psychology, cognitive science, and neuroscience, have promoted novel methods for measuring intelligence that relate directly to information processing (e.g., Jäger & Althoff, 1994; Naglieri & Das, 2005), measures used in some fields still rely heavily on intelligence-as-knowledge. In particular, industrial-organizational (I-O) psychology has undergone limited methodological developments related to intelligence in the past few decades and has rarely adopted measures that reflect an information processing approach to intelligence (e.g., Agnello, Ryan, & Yusko, 2015; Fagan &
Instead, I-O psychology researchers often rely on the tenets of Spearman’s (1927) intelligence theory, which suggest that the content of an intelligence measure is not as relevant as how strongly the measure relates to a general factor of intelligence (e.g., Dahlke & Sackett, 2017; Gottfredson, 2002; McDaniel & Kepes, 2014; Olea & Ree, 1994; Ree, Earles, & Teachout, 1994; Schmidt, 2002). These measures of general intelligence, or g, have only evolved slightly since their original development over a hundred years ago (Sternberg & Kaufman, 1996; Wasserman, 2012) and are typically rooted in measurement designs that assess intelligence-as-knowledge (e.g., Lubinski & Dawis, 1992; Roznowski, Dickter, Hong, Sawin, & Shute, 2000; Sternberg & Wagner, 1993). As a result, most measures utilized in I-O psychology are often considered heavily knowledge-based (Hunt, 2000).

The narrow use of intelligence-as-knowledge measures in I-O psychology is perpetuated by researchers emphasizing the utility of these measures (Scherbaum et al., 2012). Indeed, a great deal of importance has been placed on the effectiveness of knowledge-based intelligence tests. Validity generalization and meta-analysis studies demonstrate that intelligence measures effectively predict an array of life outcomes, including educational success, job performance, and life satisfaction (e.g., Hülsheger, Maier, & Stumpp, 2007; Hunter & Hunter, 1984; Kramer, 2009; Kuncel & Hezlett, 2007; Salgado et al., 2003; Schmidt, 2002; Ziegler, Dietl, Danay, Vogel, & Bühner, 2011). While the predictive power of these tests makes them attractive to utilize across various settings and demonstrates their utility, it has also led to a “case closed” mentality to how intelligence is measured in I-O psychology (e.g., Goldstein, Zedeck, & Goldstein, 2002; Scherbaum et al., 2012). This mentality has led researchers to rest on the collective laurels of the field, which has limited attempts to improve on the test design and
features of current intelligence measures. While the correlation between intelligence and job performance is frequently cited as .51 (Schmidt & Hunter, 1998), this finding is corrected for range restriction and reflects one particular measure of intelligence used in a structured research program. More often, research on common measures of intelligence report uncorrected correlations with job performance in the .20s (Schmitt, 2014) or low .30s (Bobko, Roth, & Potosky, 1999). This suggests that while typical measures of intelligence are important predictors of performance, there is still additional variance that can be explained by other measures. Furthermore, while intelligence measures are lauded for their predictive validity, many researchers still express concerns about the score differences that are typically seen between various groups, in particular for people of different races (e.g., Hough, Oswald, & Ployhart, 2001; Outtz, 2010; Roth, Bevier, Bobko, Switzer, & Tyler, 2001; Sackett, Schmitt, Ellingson, & Kabin, 2001; Sackett & Shen, 2009). Although some researchers argue that these score differences are a result of genetics or biological factors (Herrnstein & Murray, 1994; Jensen, 1998; Rushton & Jensen, 2005), there is growing consensus that at least some of these differences are attributable to components of cognitive tests that are associated with race, but not with actual job performance. This is referred to as performance-irrelevant race-related variance (e.g., Cole, 1973; Darlington, 1971; Newman, Hanges, & Outtz, 2007; Outtz & Newman, 2010) and suggests that there is the possibility of reducing group score differences between races on intelligence measures without reducing their validity.

Given the importance of utilizing well-validated predictors of workplace outcomes, especially due to the potentially high costs and risks associated with selection procedures (Yusko, Bellenger, Larson, Hanges, & Aiken, 2017), it is understandable why I-O psychologists favor using intelligence measures with an established utility instead of adopting intelligence
measures or measurement designs from other disciplines until there is sufficient evidence of their validity and value (Oswald & Hough, 2012). The potential impact, both positive and negative, of intelligence measures in organizations means there must be a strong, empirical vetting process prior to implementing them. However, simply choosing intelligence measures based on convention and not on an a priori evaluation of the target construct can be equally problematic as it potentially limits the domain of intelligence being assessed (Roznowski et al., 2000; Scherbaum et al., 2015). As a result, other tests of intelligence that could increase the coverage of the construct and are better aligned with the outcome of interest are not evaluated and advanced in the field, marking stagnation in how intelligence is measured and potentially leaving additional predictive power untapped.

This is particularly problematic since an applied field like I-O psychology has much to gain from leveraging information processing measures to predict important work outcomes. Organizations continue to place a premium on employees’ ability to think analytically (Aydin, Leblebici, Arslan, Kilic, & Oktem, 2005), adapt to environmental changes (Fagan & Ployhart, 2015), solve increasingly complex problems (Edmondson, 2012), and manage multiple and competing tasks (Bosco, Allen, & Singh, 2015; König & Waller, 2010), all of which are abilities associated with an information processing approach to intelligence (e.g., Gottfredson, 1997; Luria, 1973). Information processing is also believed to underlie the acquisition of new information (Ackerman, 1996), and therefore can play a critical role in how employees learn new skills and knowledge (Cowan, 2005). While there are examples of information processing measures being utilized in other applied fields, like educational psychology (e.g., Mandelman, Barbot, & Grigorenko, 2016; McCallum & Bracken, 2012; Sabet, Scherbaum, & Goldstein, 2013), and even some examples in I-O psychology (e.g., Bosco et al., 2015; Higgins, Peterson,
Pihl, & Lee, 2007), this research is still limited compared to research on intelligence-as-knowledge. By integrating additional research from the information processing approach into I-O psychology, researchers may benefit from modern and interdisciplinary theories of intelligence and capitalize on a growing number of new measures intended to evaluate one’s intelligence while reducing reliance on a person’s prior knowledge. Furthermore, this integration fits with recent calls by researchers to expand the measurement domain of intelligence in I-O psychology (Fagan & Ployhart, 2015; Reeve, Scherbaum, & Goldstein, 2015).

Measures of information processing have shown promising early evidence of predictive validity (e.g., Bosco et al., 2015; Higgins et al., 2007; Mandelman et al., 2016; McCallum & Bracken, 2012). A prior meta-analysis demonstrated that a specific type of information processing measures, fluid reasoning measures, had uncorrected correlations with job performance \( r = .14; k = 23 \), training performance \( r = .25; k = 20 \), and academic performance \( r = .22; k = .41 \), demonstrating evidence of their predictive validity (Postlethwaite, 2012). While this study provided initial evidence of the predictive power of fluid reasoning measures, it only covered a segment of the available information processing measures and it did not examine group score differences. As such, there is currently no comprehensive review of information processing measures and their predictive power. This could be a contributing factor to the lack of adoption of information processing measures by researchers in I-O psychology as they may prefer to stay “just behind the cutting edge of intelligence theories…to take advantage of theories and research findings that have already withstood a great deal of scientific and empirical scrutiny” (Oswald & Hough, 2012, p. 172). Therefore, I will conduct a meta-analysis of information processing measures to examine their ability to predict performance outcomes. Doing so will allow for a cumulative and empirical evaluation of the utility of information
processing measures for predicting performance and offer an initial test of their appropriateness for applied researchers. Furthermore, I will examine the score differences information processing measures produce between different groups, specifically between African Americans and Caucasians. Although other group score differences are equally important, group score differences between African Americans and Caucasians have been historically seen as one of the largest and most consistent differences on measures of intelligence and has led them to be a prominent point of discussion in the I-O psychology literature (Bobko, Roth, & Potosky, 1999; Newman, Jacobs, & Bartram, 2007; Sackett & Ellingson, 1997; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997). For that reason, I focus specifically on these group score differences on intelligence measures.

In the remaining sections of this dissertation, I will provide an overview of the current literature on intelligence with an emphasis on the relationship between its conceptualization and its operationalization. Then, I will offer an overview of the current state of measures that assess knowledge and information processing. Finally, I conduct a meta-analysis of information processing measures, examining their predictive validity for academic and occupational performance outcomes and examine if group score differences between African Americans and Caucasians exist on these measures.

**Intelligence**

The histories of intelligence measurement and I-O psychology have been intertwined for much of the last century. The revolutionary measurement design of Binet and Simon (1905) is heralded as the origin of modern standardized intelligence testing and has influenced many of the intelligence measures used today. This design spurred the development of the Army Alpha and Beta tests during World War I, marking the first large scale administration of intelligence tests
for selection purposes (Ackerman, 1996). Shortly thereafter, a wave of research utilizing intelligence measures in organizations established how intelligence can be applied to the workplace (e.g., Carson, 2014; Urbina, 2011). Since then, intelligence measures have been a critical part of I-O psychology, utilized by researchers to predict critical workplace outcomes and by practitioners to select new employees, amongst other uses. Individual studies and meta-analyses have demonstrated that intelligence predicts various workplace outcomes, such as job performance across various countries, occupations, organizations, and employee tenure (e.g., Bertua, Anderson, & Salgado, 2005; Levine, Spector, Menon, Narayanan, & Cannon-Bowers, 1996; Salgado et al., 2003; Schmidt & Hunter, 1998); training success (e.g., Colquitt, LePine, & Noe, 2000; Hülsheger et al., 2007; Schmidt & Hunter, 1998); counterproductive workplace behaviors (Sackett & Lievens, 2008); and prosocial behaviors (Gonzalez-Mulé, Mount, & Oh, 2014). Today, there is even more emphasis on the role of intelligence in organizations as businesses rely on knowledge and sustained innovation as competitive advantages (e.g., DeNisi, Hitt, & Jackson, 2003; Edmondson, 2012; Powell & Snellman, 2004; Scherbaum & Goldstein, 2015).

While the importance of intelligence to I-O psychology is undeniable, ambiguity surrounds what constitutes the construct. Intelligence is a latent variable and therefore must be inferred through measurement rather than be directly observed. Throughout the history of intelligence research, several traditions have arisen that differ in their methodology for inferring and understanding its latent structure. As a result, there have been numerous conceptualizations of intelligence over the years, each highlighting distinct components of intelligence, its structure, its function, and its measurement. These conceptualizations vary in their divergence from one another, making it difficult to identify a clear consensus for defining intelligence. Jensen (1998)
even argued that agreement is impossible at this point because of the diversity of definitions and
the emotionally charged nature of intelligence.

Indeed, intelligence as a construct is marred in its history, particularly its ties with
heredity and eugenics research (Wasserman, 2012) and therefore elicits many different reactions
from researchers and lay people, alike. In addition, many people have implicit theories of
intelligence that are influenced by their knowledge of intelligence measures and how these
measures are scored and utilized (Fagan, 2000). Boring (1923) famously stated that “intelligence
is what the test tests” (p. 35), a statement meant as a point of discussion rather than a conclusion
about intelligence tests, though some researchers have come to accept this as a working
definition of intelligence (Sternberg, 2003). Because many intelligence measures currently used
rely on one’s knowledge to assess intelligence, critics argue that using measurements of
intelligence to define the construct often equates intelligence to how much knowledge one has or
his/her test-taking ability (e.g., Fagan, 2000; Gardner, 1993; Sternberg et al., 2000). While
models of intelligence do include knowledge as a potential factor, today many researchers
recognize that intelligence is not just the knowledge that one has but is rather a deeper and
broader construct (Gottfredson, 1997). This has led to the development of various models that
highlight both information processing and knowledge as critical components of intelligence.

The information processing, or intelligence-as-process, approach to intelligence was
originally derived from work in the cognitive sciences that used the computer as a basis for better
understanding the human brain and highlights the architecture of cognitive processes (Sternberg,
2003). This approach assumes that people analyze information from their environment using a
set of basic cognitive mechanisms, such as perceiving, reasoning, conceptualizing, planning,
inferring, and deducing, to learn from experience and apply logical solutions to novel problems
(Sternberg, 1977, 2003). From this perspective, information processing can be conceptualized as a person’s ability to learn, maintain, and use information to reason (Sternberg, 1997). In many ways, an information processing approach to intelligence is similar to Gottfredson’s (1997) well-cited definition of intelligence as “a very general mental capacity that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly, and learn from experience” (p. 13). Furthermore, Rolfhus and Ackerman (1999) contend that these cognitive processes are best exemplified by tasks that require abstract reasoning and working memory.

This is not a new approach to conceptualizing intelligence as even Spearman (1927), one of the originators of contemporary intelligence research, referred to intelligence as a general “mental energy” rather than a specific set or sets of knowledge. Spearman (1923) further identified three principles of cognition that included perceiving and encoding the environment, inferring relationships between concepts gathered from the environment, and applying the principles inferred from those relationships to new situations. Contemporary models of information processing similarly highlight the adaptive processes involved in intelligence, demonstrating there are executive processes that organize information and allocate resources for solving problems (Lohman & Lakin, 2011; Sternberg, 1985). For example, Sternberg (1984) identified three major components of information processing, which he labeled metacomponents, performance components, and knowledge-acquisition components. Metacomponents are high-level processes that identify a problem that needs to be solved, organizes other processes required to address the problem, and evaluates progress towards accomplishing the necessary goals for solving the problem. These types of processes, sometimes referred to as executive processes, prescribe what needs to be accomplished by the performance components. The
operations executed by the performance components typically include retaining constructs in memory, manipulating constructs to identify patterns, perceiving relationships between constructs, and applying identified relationships to new stimuli (Sternberg, 1997). The knowledge-acquisition components assist with learning how to solve problems, therefore informing the performance components. For example, people will learn how to encode information that is most important or to compare features of constructs that are most relevant to the current problem (Sternberg, Kaufman, & Grigorenko, 2008). Overall, these componential models exemplify the cognitive processes seen as critical for information processing, particularly from a cognitive science perspective.

Based on the research and models being developed in the cognitive sciences, some theories have been revised to similarly suggest that intelligence involves information processing as a critical component. Modern psychometric theories, such as the Cattell-Horn-Carroll (CHC) model, depicts intelligence as a hierarchy that covers a diverse set of cognitive abilities (Schneider & McGrew, 2012; Schneider & Newman, 2015). Near the top of this hierarchy is the ability of fluid reasoning, also known as fluid intelligence or Gf, which represents a controlled set of mental operations that allow a person to solve novel problems through logic and thinking flexibly (McGrew, 2009). At the core of fluid reasoning is the idea that the type of mental operations being conducted are detached from memorization or known routines and, instead, involve inductive and deductive reasoning to draw inferences, generate and test hypotheses, and solve problems (McGrew, 2009; Wasserman & Wasserman, 2017). While fluid reasoning is a complex and multidimensional construct, all its components are seen as functional processes with the express purpose of “solving unfamiliar problems” (Schneider & McGrew, 2012, p. 111).
However, fluid reasoning often works in conjunction with acquired knowledge to solve novel problems that do not fit a person’s prior experiences (Goode & Beckmann, 2010).

Other theories have incorporated research from multiple disciplines to further the concept of information processing and its role in intelligence. Hebb (1949) introduced a rudimentary theory suggesting that information processing was connected to specific regions of the brain. Since then, additional research in neuroscience, cognitive science, and neuropsychology have rooted Hebb’s assumptions in empirical data to show that several critical cognitive processes related to information processing occur in various regions of the brain (e.g., Das, Naglieri, & Kirby, 1994; Kane & Engle, 2002). This research has been used to develop several theories suggesting these cognitive processes are involved in guiding attentional resources and solving problems, making them directly related to information processing. Included in these theories are the executive attention theory of intelligence (e.g., Heitz, Unsworth, & Engle, 2005; Redick, Calvo, Gay, & Engle, 2011) and the planning, attention-arousal, simultaneous, and successive (PASS) theory of intelligence (Das & Varnhagen, 1986; Naglieri & Das, 2005; Naglieri, Rojahn, Matto, & Aquilino, 2005). However, each theory highlights a unique set of cognitive processes, suggesting that while the theories are similar in nature, they each examine a distinct component of information processing.

It is important to highlight that these various theories of information processing do not eliminate knowledge from the realm of intelligence, but rather emphasize the interplay between the processes and knowledge of intelligence. In Ackerman’s (1996) model of intelligence-as-process and intelligence-as-knowledge, one’s ability to process information leads to the capacity to acquire knowledge from one’s experiences. This perspective, also shared by Cattell (1943), emphasizes that one’s knowledge is a result of investing information processing abilities into
learning a particular subject. Applying this logic to the workplace, this suggests that job
knowledge is a result of one’s information processing abilities interacting with job experience or
job-related information (Schmidt & Hunter, 1993). Furthermore, as one’s knowledge grows, it
can act as an input for future information processing, allowing a person to transfer knowledge to
new situations and engage in deductive reasoning based on concepts already learned (Lohman &
Lakin, 2011). Therefore, it is important to recognize that these two intelligence constructs are
inherently intertwined yet can still be conceptualized and measured independently.

From Ackerman’s model, knowledge can be conceptualized in many different forms,
such as declarative and procedural knowledge, that make up one’s knowledge structure (Rolfhus
& Ackerman, 1999). Within the CHC model of intelligence, these knowledge structures are
made up of multiple abilities, all of which are considered forms of knowledge. Perhaps most
well-known is crystallized intelligence, also referred to as comprehension-knowledge or Gc,
which entails a person’s language comprehension and general knowledge, but there is also
reading and writing (Grw), quantitative knowledge (Gq), and domain-specific knowledge (Gkn),
with each of these having additional subordinate cognitive abilities (McGrew, 2009; Schneider &
Newman, 2015). For example, included within domain-specific knowledge is various forms of
job-related knowledge, such as finance knowledge or building and construction knowledge.
However, while intelligence-as-knowledge entails domain-specific knowledge, it is most often
contemplation and operationalized broadly (Ackerman, 2000). Thus, while knowledge taken as
a whole has a fairly concrete conceptualization, it is the information a person has acquired
through experience, it is also extremely broad as it covers the facts, rules, principles, and
procedures a person knows across various domains.
While intelligence-as-knowledge and intelligence-as-process are intertwined, distinguishing between them can have significant implications for how intelligence is studied. Foremost, by conceptualizing intelligence as both process and knowledge, there are more rigid boundaries on what is and is not intelligence, reducing some ambiguity around the construct of intelligence. Furthermore, the distinction between information processing and knowledge can facilitate the advancement of theoretical and methodological approaches to intelligence. Based on recent research on information processing, prior theories of intelligence have been reassessed and integrated with theories from other disciplines to arrive at a more nuanced understanding of intelligence. For example, one of the most prominent theories of intelligence continues to be the psychometric approach. Several psychometric theories have been revitalized to reflect a more hierarchical structure that reflects both intelligence-as-process and intelligence-as-knowledge (e.g., Carroll, 1993; McGrew, 1997, 2005). Similarly, research from neuroscience has examined the link between cognitive functions and the resources of the brain (e.g., Conway, Kane, & Engle, 2003; Kane & Engle, 2002; Naglieri, 2005). Recently, research has focused on which regions of the brain are associated with knowledge, information processing, or both (Colom et al., 2009). Overall, this type of research demonstrates how information processing and knowledge complement each other and allows for a more comprehensive model of intelligence.

Conceptualizing intelligence as both a process and knowledge also offers more diverse ways to operationalize intelligence. However, it also emphasizes the importance of aligning both the conceptualization and operationalization of intelligence with the needs of a situation. This is particularly important for researchers and practitioners in organizations where intelligence is considered a critical variable for predicting outcomes. For example, an employee’s knowledge is important for job performance to the extent that he or she has the declarative (i.e., job
knowledge) and procedural knowledge (i.e., job skills) needed to perform work tasks (Kanfer & Ackerman, 2005). Therefore, an intelligence-as-knowledge approach is particularly useful when understanding and measuring an employee’s job relevant knowledge (e.g., measuring what an employee might need to know on day one of a new job). Extant research demonstrates that both measures of specific knowledges (e.g., quantitative knowledge) (Goertz, Hülsheger, & Maier, 2014; Ziegler et al., 2011) and global measures of one’s knowledge (i.e., not job specific knowledge) (e.g., Hunter & Hunter, 1984) predict performance at work. To date, many of the studies and meta-analyses examining the link between intelligence and performance have examined global measures of intelligence-as-knowledge, such as the Wonderlic Test (e.g., Hunter, 1989; Jansen et al., 2013) the General Aptitude Test Battery (e.g., Hunter & Hunter, 1984), and the Armed Services Vocational Aptitude Battery (e.g., Ree & Earles, 1991; Roberts et al., 2000).

In contrast, information processing is critical for learning information, including new knowledge and skills (Kanfer & Ackerman, 2005). Therefore, information processing is likely a key predictor of employee success in contexts that require handling novel situations or the acquisition of new knowledge or skills (Cowan, 2005). As a result, while information processing does not represent the actual knowledge or skills an employee has, it represents at least one predictor of an employee’s potential to gain new knowledge and learn new competencies or skills. In addition, information processing is believed to be particularly useful for understanding job performance on complex jobs or those that require high levels of attentional resources (e.g., financial analysts) (Kanfer & Ackerman, 2005). For these reasons, information processing measures rely less on a person’s prior knowledge and more on the broad processes involved in reasoning, solving problems, and/or learning. However, regardless of how intelligence is
operationalized, it is best treated not as a singular construct, but rather a network of interrelated constructs that offer insight into the various manifestations of intelligence, mainly as a process and as content (Reeve et al., 2015).

Both intelligence-as-knowledge and intelligence-as-process clearly have implications for understanding workplace behaviors and have the potential to complement each other, as Ackerman (1996) originally proposed. Specifically, it is important to align the needs of the testing context with the content of the test being employed. Measures of broad knowledge, such as crystallized intelligence, can be particularly useful when the criterion requires a person to have a large and broad knowledge base. Alternatively, in situations where one’s ability to process novel information efficiently and effectively is of interest, information processing measures may provide a better prediction of the criterion. Such an approach suggests that evaluating the correspondence of the criterion and the predictor are of upmost importance (Judge & Kammeyer-Mueller, 2012). In the case of many jobs, it is feasible that both knowledge and information processing are critical to effective performance (Schneider & Newman, 2015).

However, while the majority of research in I-O psychology has examined the effect of knowledge on critical performance outcomes, there has yet to be an extensive review of the effects of information processing on performance outcomes. In the subsequent sections, I will provide an overview of the research and measures associated with both intelligence-as-knowledge and intelligence-as-process to offer additional insight into the gap that currently exists in understanding information processing at work.

**Measuring Intelligence-as-Knowledge**

Traditionally, intelligence measures in I-O psychology have been heavily focused on a person’s knowledge. Fitting with the conceptualization of intelligence-as-knowledge, the current
study focuses on knowledge-based measures, such as crystalized intelligence measures, that assess broad knowledge rather than domain-specific knowledge. Though there are intelligence-as-knowledge measures that assess job-specific knowledge, many measures assess a person’s general knowledge. Hunt (2000) speculated that most measures used in applied fields focus on intelligence-as-knowledge, specifically crystallized knowledge. Similarly, researchers have evaluated the content of intelligence measures based on a $Gf$-$Gc$ framework and concluded that most measures assessed crystallized intelligence and not fluid reasoning (Alfonso, Flanagan, & Radwan, 2005; McGrew, 1997). The heavy reliance on knowledge measures, particularly general knowledge measures, has mainly resulted from the theoretical foundations that underlie much of the research in I-O psychology. Specifically, the psychometric approach to intelligence has been the focal theory of I-O psychology for decades and has influenced how the field measures intelligence.

The psychometric approach to intelligence relies on the use of measurement and statistical analyses to identify the underlying structure of intelligence (Embretson & McCollam, 2000). Many of the intelligence measures used over the past century came out of the psychometric approach (Buros, 1977; Carroll, 1978; Scherbaum et al., 2012). Spearman (1927) originated the psychometric approach by relying on the use of factor analysis, a methodology he pioneered, to understand the latent source of individual differences on test performance. He found that scores on distinct intelligence tests were positively intercorrelated with one another, which he referred to as positive manifold. Spearman and proponents of his model used this evidence to suggest that there is a general factor of intelligence, known as $g$, which explains the shared variance across intelligence measures. Furthermore, based on the idea of positive manifold, proponents of the psychometric approach argue that any measure of intelligence that
correlates highly with $g$ will be a good measure of intelligence, regardless of content (Gottfredson, 2002; Ree, Carretta, & Teachout, 2015; Spearman, 1927). This phenomenon is known as the *indifference of the indicator* and suggests that $g$ can be measured across a wide range of methodologies and measures (Spearman, 1927). This has important implications for test design since a priori theory about the content of a test does not matter as much as its loading onto a single factor (which is known as a $g$-loading). Jensen (1998) contended that a test battery of diverse, but highly $g$-loaded measures should remove unwanted variance and aggregate to create an accurate measure of $g$. From this perspective, the content of an intelligence measure matters less than the psychometric properties of the measure (Jensen, 1992). Thus, according to proponents of Spearman’s psychometric approach, most intelligence measures are interchangeable and will serve as a measure of $g$. While there have been measures of information processing to come out of the psychometric approach (as discussed in the following sections), most psychometric intelligence measures are knowledge-based (McGrew, 1997).

There are many intelligence tests grounded in Spearman’s psychometric approach that are composed of knowledge items and employed in organizational contexts. Some of the most well-known intelligence tests in the workplace, including the Wonderlic Test and the Armed Services Vocational Aptitude Battery, consistently correlate with crystallized intelligence, but not with fluid reasoning (Bell, Matthews, Lassister, & Leverett, 2002; Hicks, Harrison, & Engle, 2015; Matthews & Lassiter, 2007; Roberts et al., 2000). These measures have been shown to predict performance in both work and school (Barrick, Mount, & Strauss, 1993; K. G. Brown, Le, & Schmidt, 2006; Chamorro-Premuzic & Furnham, 2008; Frei & McDaniel, 1998; Knapp, Campbell, Borman, Pulakos, & Hanson, 2001; McKelvie, 1994). Similarly, many of the frequently cited meta-analyses demonstrating the relationship between intelligence and job
performance (e.g., Hunter & Hunter, 1984; McDaniel, Schmidt, & Hunter, 1988; Schmidt & Hunter, 2004; Schmidt & Hunter, 1998) rely heavily on the General Aptitude Test Battery, which consists of assessments of one’s knowledge of vocabulary and computation (te Nijenhuis & van der Flier, 1997). Even measures that are intended to assess a general form of intelligence (i.e., not described as a measure of knowledge) are typically heavily weighted with knowledge subtests and dimensions (Schneider & Newman, 2015). Some argue that these knowledge-based measures act as proxies of a person’s ability to learn (e.g., Hunter, 1986; Hunter & Schmidt, 1996; Schmidt, 2002), however they remain infused with items that require prior knowledge for test takers to answer correctly. Therefore, while these measures may assess parts of information processing, such as learning ability, the resulting scores are not a pure assessment of information processing. Regardless, these measures have been shown to consistently predict critical workplace and educational outcomes, engraining their use in the I-O psychology field. The high utility of these measures, in combination with the principle of the indifference of the indicator, has likely perpetuated the reliance of applied psychologists on knowledge-based measures.

While measures of knowledge have been praised for their ability to predict workplace and educational outcomes, they have been critiqued on several fronts. Foremost is the concern that tests relying solely on knowledge-as-intelligence are too narrow in their content (Agnello et al., 2015; Gottfredson, 1997; Scherbaum et al., 2012). Knowledge-based measures of intelligence only capture a portion of the intelligence construct and therefore offer a limited perspective (e.g., Alfonso et al., 2005; Chen & Gardner, 2012). While these types of knowledge measures can and do explain a considerable portion of variance in outcome variables, their narrow scope restricts them from predicting additional variance. Performance dimensions are often multidimensional, covering a wide scope of behaviors and outcomes. Therefore, using a
restricted conceptualization and operationalization of intelligence will similarly restrict the amount of variance of complex outcome that can be predicted (Higgins et al., 2007; Scherbaum et al., 2012). As an example, many of the measures that have been designed from the knowledge-as-intelligence paradigm treat problems as well-defined and therefore do not effectively measure a person’s ability to handle novel and abstract problems (Pretz, Naples, & Sternberg, 2003). Therefore, this presents a potential deficiency in the measurement of intelligence. This coincides with a second concern, mainly that the development and application of knowledge measures often rely heavily on the statistical properties of the measure and less on a theoretical foundation (Fagan, 2000; Lievens & Reeve, 2012; Reeve et al., 2015; Scherbaum et al., 2015). Recently there has been an improvement in the number of intelligence measures that have been rooted in theory, however, this trend has not fully carried over to I-O psychology (Scherbaum et al., 2012; Thorndike, 1997). To date, there has been little integration of theoretical models from other disciplines into the application of intelligence measures in I-O psychology, leaving an opportunity for additional advancements in intelligence measurement.

A third concern facing knowledge measures is their tendency to exhibit group score differences, particularly between races (e.g., Aguinis & Smith, 2007; De Corte, 1999; Goldstein et al., 2002; Ployhart & Holtz, 2008; Sackett et al., 2001; Sternberg & Wagner, 1993). Many knowledge tests have been found to favor Caucasian and Asian test-takers over African American test-takers (Hough et al., 2001; Neisser et al., 1996; Roth et al., 2001). Researchers argue these differences can be attributed to socioeconomic differences in education or cultural differences (e.g., Cottrell, Newman, & Roisman, 2015; Fagan & Holland, 2007). When these intelligence measures are utilized to make hiring decisions, these group score differences can lead to a disparity in the selection rate of people of a certain race, also known as adverse impact
Findings show that removing knowledge that varies by race from a test reduces group mean differences (Fagan & Holland, 2007; Freedle & Kostin, 1997; Malda, van de Vijver, & Temane, 2010), and the size of group mean differences often depends on the type of test used (Hough et al., 2001). Thus, there seems to be support for the notion that group score differences are due, at least in part, to the content of these measures, specifically the use of knowledge items, and is not solely a result of a person’s intellectual ability.

Taken together, these critiques suggest that while knowledge measures can reflect important content on a job and have high face validity (Lang, Kersting, Hülsheger, & Lang, 2010; McDaniel & Banks, 2010), they only assess one aspect of intelligence (Roznowski et al., 2000). Therefore, despite their predictive power, researchers should question the reliance of I-O psychology on knowledge measures of intelligence, identifying areas where a more inclusive approach to the conceptualization and operationalization of intelligence can be beneficial and may reduce group score differences on intelligence measures (e.g., Agnello et al., 2015; Fagan & Ployhart, 2015; Scherbaum et al., 2015; Scherbaum et al., 2012). To date, there has been a lack of exploration into other methodologies for measuring intelligence by applied psychologists, such as the intelligence-as-process perspective.

**Measuring Intelligence-as-Process**

Researchers have proposed that intelligence measures place too much emphasis on knowledge since the outset of modern intelligence measurement (Cattell & Bristol, 1933). In recent years, researchers have noted that most test batteries do not include adequate measures of a person’s reasoning ability (Alfonso et al., 2005; McGrew, 1997, 2005). This is unfortunate, as information processing measures can address calls for additional measures to help capture a
broader scope of the intelligence construct (e.g., Bosco et al., 2015; Cowan, Fristoe, Elliott, Brunner, & Saults, 2006; Fagan & Ployhart, 2015; Roznowski et al., 2000). In fact, even within the intelligence-as-process framework there is a wide range of cognitive constructs that constitute information processing, such as fluid reasoning, working memory, and executive function (Krumm, Lipnevich, Schmidt-Atzert, & Bühner, 2012; Krumm et al., 2009). While these constructs are related, they are not considered isomorphic and, as a result, elucidate various cognitive processes associated with information processing (e.g., Ackerman, Beier, & Boyle, 2005; Clancy Blair, 2006; Oberauer, Schulze, Wilhelm, & Süß, 2005). These constructs are derived from different theoretical traditions and result in various measures for assessing specific aspects of information processing. For the purposes of this paper, I will focus on two overarching categories for understanding and measuring information processing: 1) the psychometric approach and 2) the cognitive science approach.

The psychometric approach. Since Spearman’s traditional psychometric approach to intelligence, which advocated for a single factor of intelligence, significant evolutions have reframed intelligence as hierarchical. In one of the earliest hierarchical models of intelligence, Cattell and Horn (Cattell, 1963, 1971; Horn & Cattell, 1966) argued for the distinction between fluid reasoning and crystallized intelligence. While psychometric models have further evolved, fluid reasoning continues to be a critical part of the psychometric approach, including in the dominant contemporary model, the CHC model (Schneider & McGrew, 2012; Schneider & Newman, 2015). In this multidimensional approach, fluid reasoning refers to a person’s ability to solve novel problems through abstract reasoning and identifying patterns and subsumes other cognitive abilities, such as deductive reasoning, inductive reasoning, quantitative reasoning, and
speed of reasoning (McGrew, 2009). Within the psychometric approach, fluid reasoning is clearly tied to information processing.

The CHC model also contains other factors that may be considered aspects of information processing. However, these factors either have their theoretical roots in the cognitive sciences approach or have imprecise ties to intelligence. Two examples of these factors include working memory and processing speed. For the past several decades, research on working memory has been clearly established in the cognitive science approach as researchers have examined the critical brain regions associated with working memory (e.g., Cohen et al., 1997; D'Esposito & Postle, 2015). More recently, working memory has been integrated into the psychometric approach through factor analysis findings (Carroll, 1993; McGrew, 2009) and by examining the construct’s relationship with fluid reasoning (e.g., Conway et al., 2003; Kyllonen & Christal, 1990). While research on working memory from the psychometric approach demonstrates the benefits that can come from incorporating research from other disciplines, such as the cognitive sciences, the theory behind the working memory factor remains rooted in the cognitive science approach and therefore will be addressed later in this paper.

Another factor in the CHC model that has ties to information processing is processing speed. While processing speed has been examined in relation to constructs such as fluid reasoning (e.g., Conway et al., 2003; Fry & Hale, 1996), its overall relationship with intelligence is muddled. Researchers have suggested that the meaning of processing speed differs based on the operationalization of the construct (Carroll, 1993; Conway et al., 2003). Specifically, measures involving reaction times based on simple stimuli and those based on complex stimuli likely stimulate different cognitive processes. This means that when processing speed measures are taken as a whole, their relationship with intelligence is not entirely clear (Neisser et al.,
Along these lines, Carroll (1993) suggested that some measures of processing speed do not reflect a person’s capacity to process complex information, but rather simply how quickly they can respond to a stimulus. Carroll further argued from this perspective that some measures of processing speed are not a sign of intelligence but rather a sign of speediness in achieving an outcome.

Researchers in the cognitive sciences have begun to parse out the different types of processing speed measures and utilize those that involve complex stimuli and reactions. In these instances, the measures are being associated with high-level cognitive processes, such as attention, usually within a battery of other measures (e.g., Bosco et al., 2015; Bühner, König, Pick, & Krumm, 2006; Higgins et al., 2007; Sabet et al., 2013). This subset of processing speed measures seems to better assess a person’s information processing ability than those that require reactions to simple stimuli. However, while these measures are similar or the same to those being used to assess processing speed, they are more clearly rooted in the theory of the cognitive sciences approach. For these reasons, I do not include measures of processing speed as part of the psychometric approach as they either do not assess information processing as conceptualized in this paper or, when they do measure information processing, are linked to the cognitive science approach.

Therefore, while the psychometric approach has begun to implement additional measures of information processing into its contemporary models of intelligence, this research remains nascent and needs further integration to fully understand its relationship to psychometric models. To date, the most deeply engrained construct of information processing from the psychometric approach is fluid reasoning. For the purposes of this paper, I will exclusively examine measures
of fluid reasoning as representative of information processing measures from the psychometric approach.

There have been multiple fluid reasoning measures developed over the years, though they mostly share the same intent of reducing the need of prior knowledge (Cattell, 1987). The presence of items that rely heavily on prior knowledge can reduce the construct validity of the measure by introducing unrelated content and provide an advantage to those familiar with that content, causing misleading score differences. Specifically, research has shown that knowledge of content unrelated to information processing may differ by people’s background, country of origin, race, gender, culture, economic standing, or experience, leading to observed score differences (Fagan, 2000; Fagan & Holland, 2002; Goldstein et al., 2010). Similarly, research shows that knowledge of culturally specific information (Malda et al., 2010) or exposure to acquired knowledge (Fagan & Holland, 2002, 2007, 2009) on an intelligence test can moderate the relationship between race and observed scores (Malda et al., 2010). Therefore, to avoid construct irrelevant variance within fluid reasoning scores and, subsequently, reduce the accuracy of test scores, test developers often utilize non-entrenched task designs (e.g., Bokhorst, 1989; Tetewsky & Sternberg, 1986). Non-entrenched tasks require information processing strategies that are outside a person’s normal experience, meaning they do not rely on automatized and/or previously learned performance routines. Typically, these tasks include stimuli that are novel to participants and do not directly reflect how a problem and its solution would appear in everyday life (Sternberg, 1982). For example, by using stimuli such as unfamiliar images (i.e., non-verbal stimuli) or concocted words (i.e., pseudo-words), the test taker must identify the relationship between stimuli to find a solution. Through this type of task, all test takers are placed on an even playing field, making it possible to assess their true information
processing ability. This approach helps ensure the measures assess the core competencies of fluid reasoning, mainly a person’s ability to reason to solve novel problems and minimize the reliance on a test taker’s knowledge (e.g., Fagan, 2000; Fagan & Holland, 2002, 2007, 2009; Helms-Lorenz, Van de Vijver, & Poortinga, 2003; Malda et al., 2010; van de Vijver, 1997). Since most measures of fluid reasoning utilize these types of non-entrenched tasks, the main difference between measures is in the format of the task, specifically whether it uses a graphical strategy or a verbal strategy for presenting stimuli.

**Graphical tests.** Perhaps the most prototypical type of fluid reasoning measures is those that utilize graphical stimuli, particularly geometric matrix items (Marshalek, Lohman, & Snow, 1983; Primi, 2014). In general, these measures rely on various graphical images to present stimuli for the test taker to infer relationships. To exemplify the structure of these types of measures, I discuss the Raven’s Progressive Matrices and the Cattell Culture Fair Test, two of the oldest fluid reasoning measures still in use.

**Raven’s Progressive Matrices.** One of the most frequently utilized measures of fluid reasoning is the Raven’s Progressive Matrices (Urbina, 2011; van de Vijver, 1997). The Progressive Matrices is comprised of a series of items that contain a matrix (e.g., a 3 x 3 matrix). Each cell of the matrix contains a figure with a unique set of shapes and lines except one blank cell (Raven, Raven, & Court, 2000). The test taker must review the available figures and identify the pattern between them. By determining the pattern, the test taker must then apply his/her logic to select the figure that belongs in the empty cell from a set of alternatives. There have been several versions of the Progressive Matrices available, including the Standard Progressive Matrices, which is the original version of the test published in 1938, the Colored Progressive Matrices, which is designed for children between 5 and 11 years old, and the Advanced
Progressive Matrices, which are a more difficult version of the Standard Progressive Matrices and most appropriate for adults. While the test typically contains 60 items, abbreviated versions of the test have also been developed (Bilker et al., 2012). Several researchers have proposed that the Progressive Matrices are a particularly good assessment of fluid reasoning because it requires control processes to analyze the problem, strategize how to approach it, and apply a solution (e.g., Carpenter, Just, & Shell, 1990; Marshalek et al., 1983).

**Cattell Culture Fair Intelligence Test.** Much like the Progressive Matrices, the Cattell Culture Fair Intelligence Test (Cattell & Cattell, 1973) is specifically intended to reduce the effect of previous or cultural knowledge through its reliance on graphical stimuli. There are four subtests to the Culture Fair Test, including a matrices subtest, which mirrors the items in the Progressive Matrices. There is also a series subtest, which is similar to the matrices, but requires the test taker to find the pattern across a series of matrices rather within a single matrix, and then select a matrix to complete the series. In addition, there is a classification subtest in which the test taker must identify the odd figure or figures in a set. Finally, there is the topology subtest, which requires test takers to select a figure based on a set of constraints described in the instructions (e.g., choose the figure where the dot is in the circle but outside the square). These subtests have been shown to correlate well with each other and likely assess a person’s inductive reasoning skills by requiring them to identify both similarities and differences between stimuli (Troche, Wagner, Schweizer, & Rammsayer, 2016).

**Verbal tests.** Verbal measures of fluid reasoning offer a unique approach for assessing reasoning. Research has demonstrated that verbal stimuli elicit unique cognitive processes (Baddeley, 2012), making verbal measures of fluid reasoning a critical area of research to better understand how reasoning manifests in different contexts. A main strategy for balancing the need
to minimize prior knowledge with items that contain language is for the measure to contain simple words that are familiar to most test takers (Cattell, 1987). Perhaps the most prominent exemplar of this strategy is the Baddeley Reasoning Test.

*Baddeley Reasoning Test.* The Baddeley Reasoning Test (Baddeley, 1968) was developed with the intention of creating a short and easily administered measure of intelligence. Baddeley stated that the measure utilizes variation in syntactic structure to fluctuate item difficulty while using language that is familiar to most test takers so that it can quickly assesses information processing. The measure consists of 64 items that are administered in three minutes. The test taker is presented with various form of statements including a rule (e.g., A precedes B) and then a statement (AB), to which the test taker must determine if the statement is true or false. Rules vary be whether they are positive or negative, active or passive, true or false, involve proceeding or following, and if A or B comes first. In addition, statements vary in whether they are AB or BA. Therefore, while the measure includes verbal stimuli, the items do not rely on extensive knowledge of language to be completed. Research has since used the Baddeley Reasoning Test to evaluate the relationship between fluid reasoning and performance, particularly in educational settings (Chamorro-Premuzic & Furnham, 2008; Chamorro-Premuzic, Quiroga, & Colom, 2009; Furnham, 2012).

*Psychometric measures and performance.* Fluid reasoning measures remain the leading strategy for assessing information processing and have become a strategy for hiring intelligent employees because of their ease of administration and purported reduction in group score differences (e.g., Klein, Pohl, & Ndagijimana, 2007; Taylor, 2008). The exact relationships between fluid reasoning measures and various performance metrics at work and school have interested researchers for decades since these measures do not require extensive prior knowledge.
to assess information processing. Several decades of research have shown fairly consistent relationships between fluid reasoning measures and job performance (e.g., Côté & Miners, 2006; Durso, Bleckley, & Dattel, 2006; Henderson, 1979; Lowery, Beadles Li, & Krilowicz, 2004), job-oriented training performance (e.g., de Bruin, de Bruin, Derksen, & Cilliers-Hartslief, 2005; Moran, 1986), and academic performance (e.g., Chamorro-Premuzic & Furnham, 2008; Chamorro-Premuzic et al., 2009; Furnham, 2012; Furnham, Zhang, & Chamorro-Premuzic, 2006; Higgins et al., 2007). Research has even found that fluid reasoning predicts specific outcomes, such as annual salary (Colonia-Willner, 1998), and performance on specific tasks, such as multi-tasking (Bühner et al., 2006) and proofreading (Furnham, Rawles, & Iqbal, 2006). Taken as a whole, these measures offer strong evidence of their predictive validity for performance outcomes.

**Psychometric measures and group score differences.** Fluid reasoning measures are typically constructed to minimize reliance on prior knowledge through the use of non-entrenched items and minimal language. This has led many fluid reasoning measures to be labeled as “culture fair” tests as they are intended to reduce the impact of a person’s cultural background on test performance (e.g., Cattell, 1971; Klein et al., 2007; Nenty & Dinero, 1981). Past research has shown that these types of item features lead to small group score differences (e.g., Fagan & Holland, 2007; Sternberg, 1986). As such, it is often expected that fluid reasoning measures, based on their item content, should have low score differences between groups, particularly between African Americans and Caucasians. Research has provided evidence to support this notion, showing various operationalizations of fluid reasoning to have effect sizes considered small to moderate (e.g., Hough et al., 2001; Kirkpatrick, Ewen, Barrett, & Katzell, 1968; Outtz & Newman, 2010). However, other research has also demonstrated that common fluid reasoning
measures, such as the Raven’s Matrices, can have fairly large score differences similar to common measures of crystallized intelligence (e.g., Bosco et al., 2015) and as large as one standard deviation (e.g., Jensen, 1998; Rushton & Skuy, 2000). Researchers have begun to evaluate why these group score differences persist under certain circumstances. For example, researchers have found larger group score differences when the testing situation or the content of the test cause threatening social stereotypes to become salient (R. P. Brown & Day, 2006; Klein et al., 2007). Others suggest that differences in exposure to technology, games, and learning materials that are similar to the tasks on fluid reasoning tests can lead to group score differences (Greenfield, 1998). While these explanations offer some initial insight, more research is needed to better understand the specific features of fluid reasoning tests that lead to variation in score differences.

**Summary of the psychometric approach.** To date, the research on fluid reasoning measures has demonstrated promising results regarding their ability to predict performance in several domains and evidence suggests that these measures may reduce group score differences, though not as much as once thought. It is important to note that much of the research to date has been conducted on the Progressive Matrices and Cattell’s Culture Fair Test, though many other measures exist. While these measures follow similar test designs as described above, there is room for further exploration of test designs and item content within fluid reasoning measures to evaluate how they influence predictive validity and group score differences. Furthermore, recent attempts have begun to examine how intelligence measures from the psychometric approach converge with findings from other fields, particularly those from the cognitive sciences (e.g., Ackerman et al., 2005; Bosco et al., 2015; Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008; Colom et al., 2009; Kane, Hambrick, & Conway, 2005). Researchers have begun to
integrate test designs from these disciplines to improve the measurement of fluid reasoning (e.g., Primi, 2014), though the predictive validity of these measures has yet to be tested widely. It is important for researchers in the psychometric approach to continue incorporating theoretical findings from other disciplines to optimize their measures (e.g., Reeve & Hakel, 2002; Scherbaum et al., 2012; Tenopyr, 2002).

**The cognitive science approach.** While the psychometric approach to intelligence has a long and rooted history in applied psychology, acting as the dominant approach for many years, research in cognitive science is gaining momentum (Becker, Volk, & Ward, 2015). Whereas the psychometric approach emphasizes factor analysis to construe the structure of intelligence, approaches involving cognitive science apply interdisciplinary theories of the mind and methods originating from neuroscience to examine how people guide and direct cognitive resources to solve problems and obtain goals (Becker et al., 2015; Drasgow, 2013). Cognitive science is comprised of multiple branches, including cognitive psychology, which examines executive functions of the brain relative to intelligence (e.g., Miyake, Friedman, Emerson, Witzki, & Howerter, 2000), neuropsychology, which examines the relationship between the brain and behavior (e.g., Conway et al., 2003; Kane & Engle, 2002; Meier, 1974), neurocognition, which links cognitive resources to specific regions of the brain through neuroimaging techniques (e.g., Becker et al., 2015; Curtis & D'Esposito, 2003; Deary, Penke, & Johnson, 2010; Jung & Haier, 2007; Nisbett et al., 2012), and more. The various functions within cognitive science supplement each other, offering deeper insights into the functioning of the brain and how it relates to the management of cognitive resources and the processing of information. For example, neuropsychology examines how people’s cognitive functions connect to behavior, which is validated through neuroimaging or other neuroscience techniques. As such, each branch offers a
theory-driven approach to understand how individual differences in intelligence are directly tied to functioning in the brain. To date, research in the workplace has mostly relied on cognitive psychology and neuropsychology, though there are future opportunities to apply additional branches from the cognitive sciences.

In the following sections, I review theory and measures associated with the two main disciplines within the cognitive sciences being applied to workplace and academic settings, cognitive psychology and neuropsychology. Because the theories and measures are rapidly evolving, this review is not meant to be comprehensive but rather offer a representative sample of those that are at a state of development that can be practically applied in organizational or educational settings. As such, while the cognitive sciences are potentially best known for their use of physiological measures, such as neuroimaging and eye tracking, I do not discuss them here, though they offer an interesting area for future research.

**Cognitive psychology measures.** Intelligence research in cognitive psychology has focused on understanding the cognitive processes of information processing and how they are organized (Sternberg, 2003). While there is no agreement in cognitive psychology about what mechanisms explain behavior, let alone intelligence, there are several theoretical models of information processing that emphasize similar processes. Typically, these models depict information processing as a system involving cognitive mechanisms associated with holding information in memory, directing attention, and controlling action (e.g., Baddeley & Hitch, 1974; Cowan, 1988; Engle, Tuholski, Laughlin, & Conway, 1999). Initially, much of the research in the field examined short-term memory, long-term memory, and attentional control as distinct constructs with their own unique contribution to cognition. Eventually these cognitive processes were assembled together in the theory of working memory, though they remain distinct today.
Through this evolution, the theory of memory and its measurement has become increasingly rooted in physiology (e.g., Engle et al., 1999; Kane et al., 2005). At its core, working memory describes how people retain diverse pieces of information in memory for short periods of time, maintain attention on relevant information, and integrate those pieces of information (Baddeley & Hitch, 1974; Prabhakaran, Narayan, Zhao, & Gabrieli, 2000). From this perspective, working memory is critical to a person’s ability to reason with information as it provides the foundational capacity to compare different pieces of information to arrive at conclusions (Kent, 2017). Researchers have even questioned if working memory is the neuropsychological form of fluid reasoning (Kyllonen & Christal, 1990), though findings show that while the constructs are highly related and activate similar regions of the brain, they are not isomorphic (Engle et al., 1999; Kane et al., 2005).

Embedded within the theory of working memory is the construct of executive attention, an underlying process that specifically coordinates what is attended to in working memory. While some researchers have questioned whether working memory and executive attention are the same construct (Engle, 2018), others have treated executive attention as an underlying process of working memory and other high-level cognitive processes (e.g., Clancy Blair, 2006; Conway et al., 2005; Duncan, Emslie, Williams, Johnson, & Freer, 1996; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). Compounded by the fact that there is no consensus on the definition of working memory (McCabe et al., 2010), the boundary between working memory and executive attention can be difficult to delineate. While recent research has distinguished slight differences in how working memory and executive attention are measured, even the measurement strategies for these constructs overlap significantly, containing some of the same measures (Bosco et al., 2015). Both constructs have been used in the literature to assess
information processing and have demonstrated earlier evidence of their ability to predict performance in multiple settings (e.g., Ackerman & Beier, 2007; Bosco et al., 2015; Bühner et al., 2006; Perlow, Jattuso, & Moore, 1997; Roznowski et al., 2000). Because it remains unclear how deeply working memory and executive attention are intertwined, I discuss these constructs together, providing an overview of both constructs and how they relate to each other conceptually and operationally.

**Working memory and executive attention.** In its most basic form, working memory refers to a limited capacity memory system in which information is held and manipulated (Baddeley & Hitch, 1974; Baddeley, 2012). Expanding on this conceptualization, researchers consider there to be two main components of working memory: 1) a storage component and 2) an attentional and maintenance component. Though the storage component originated with research on short-term and long-term memory, it has come to be differentiated by its inclusion of information storage systems specific to phonological and visual based modalities. The determination of the information to be stored, and the manipulation of that information, is completed by the attentional and maintenance component, also known as executive attention. Attention control is an important construct in cognitive psychology, predating the theory of working memory, but has evolved significantly from that time. In its simplest form, executive attention is a central executive component that involves directing attention to ongoing cognitive processes based on a person’s current needs (Engle et al., 1999). This basic concept was integrated into the theory of working memory, which specifically pointed to executive attention as the mechanism that coordinates and maintains information held in storage (Baddeley & Hitch, 1974). However, over time, executive attention has been applied to other cognitive functions (e.g., Barrett, Tugade, & Engle, 2004; McCabe et al., 2010).
Researchers have further divided working memory into three components – updating, maintenance, and inhibition (Redick et al., 2011) – though these components have been associated more specifically with the executive attention process (Bosco et al., 2015; Ren, Altmeyer, Reiss, & Schweizer, 2013). The updating function involves monitoring working memory and adding or removing information when necessary. This ensures that the information being stored in working memory is the most relevant to the task at hand. Maintenance is a related process in which information is retrieved from long-term memory or brought into working memory (Unsworth & Engle, 2007). The final process, inhibition, is a person’s ability to remain focused on the information at hand and not be distracted by potentially interfering information. Taken together, this suggests that working memory and executive attention encompasses the deliberate cognitive action to maintain attention on information relevant to one’s goals, to inhibit irrelevant information to one’s goals, and to ultimately direct appropriate responses to obtain one’s goals (Heitz et al., 2005; Redick et al., 2011).

Working memory measures build on these basic tenets, requiring test takers to complete tasks that require updating, maintenance, and inhibition. While there are many types of working memory tasks available (e.g., n-back tasks, visual array comparison tasks) (Becker et al., 2015), to date, complex span tasks are the most frequently applied to predict performance. Complex span tasks require test takers to perform various simple cognitive tasks, including reading, calculations, and determining whether a matrix is symmetrical (Case, Kurland, & Goldberg, 1982; Daneman & Carpenter, 1980; Turner & Engle, 1989). However, between these tasks, test takers must also remember distinct information to be recalled later, such as letter, digits, or words. For example, in an operation span task (OPSAN), the test taker is presented with an item that must be recalled later and then completes between two and seven simple math problems.
The test taker is then shown another item to be recalled later and then completes additional math problems (Turner & Eagle, 1989). The final score for the test involves how many items are recalled correctly. In this process, the test taker must maintain a list of items to be recalled and update the list of recall items as the process continues, all while inhibiting the distraction of the math problems. Therefore, requiring test takers to alternate between tasks while simultaneously storing information is intended to assess the key functions of working memory (McCabe et al., 2010).

A key aspect of working memory measures is that they assess both storage and executive attention processes. To assess executive attention as a distinct construct, researchers have recently suggested using a mix of working memory scales as well as simple attention tasks (Bosco et al., 2015; Hutchison, 2007). In doing so, these measures are intended to capture the processes that overlap between working memory and executive attention, as well as presumably assess the basic attention processes unique to executive attention. As an example, researchers have utilized the arrow flanker task in combination with complex span tasks for assessing executive attention. The arrow flanker task presents test takers with a series of arrows that are all congruent (i.e., pointing the same direction) or has the center arrow incongruent (i.e., pointing the opposite direction of the other arrows; Eriksen & Eriksen, 1974). The test taker must quickly and accurately determine if the series is congruent or not. Therefore, this process tests inhibition to the extent the test taker can ignore distracting stimuli and does not rely on the storage component of working memory.

*Cognitive psychology measures and performance.* To date, research from the cognitive psychology approach has relied heavily on the use of complex span tasks of working memory. Evidence from this research has demonstrated that these measures relate to performance on job
sample tests (Bergersen & Gustafsson, 2011), overall job performance (Henderson, 1979; Nelson, 2003), skill acquisition (Perlow et al., 1997), and academic performance (Rohde & Thompson, 2007). Furthermore, these measures have been found to correlate with performance on specific tasks, such as multi-tasking (Bühner et al., 2006) and a flight simulator (Van Benthem & Herdman, 2016). Additional research on executive attention has similarly shown these measures predict job performance and success on a simulation of managerial ability (Bosco et al., 2015). As such, there is a growing set of findings supporting the use of these cognitive psychology measures to predict performance.

Cognitive psychology measures and group score differences. There is currently little data examining group score differences between African Americans and Caucasians for cognitive psychology measures in I-O psychology. In the few studies that have examined group score differences, results have demonstrated promising reductions in differences. Bosco and colleagues (2015) utilized a composite of multiple complex span measures and the arrow flanker task to assess executive attention. They found across four studies that Caucasians scored a half standard deviation higher than African Americans on the executive attention composite. This effect size is considered moderate in size but was significantly smaller than the one standard deviation difference found between the groups on the Wonderlic, a measure mainly consisting of crystallized intelligence. Research from an unpublished dissertation in I-O psychology showed a similar finding in which two working memory scales had a much lower group score difference than a measure of general intelligence and even a measure of fluid reasoning (Nelson, 2003). While the research on group score differences remains fairly lacking for cognitive psychology measures, the early evidence is promising.
Summary of cognitive psychology measures. Overall, working memory and executive attention measures offer a unique measurement approach, assessing multiple cognitive mechanisms (e.g., storage, updating, maintenance, inhibition) involved in general information processing. Therefore, these measures could be extremely useful in organizations that require memorization and visual comparisons (Becker et al., 2015) and have been found to predict critical workplace behaviors. While the theory of working memory originated from theoretical models associating cognitive processes to parallels in computer processing (Sternberg, 2003), recent neurological research has based these and similar processes in regions of the brain (D’Esposito, Ballard, Zarahn, & Aguirre, 2000; McIntosh, Grady, Haxby, Ungerrleider, & Horwitz, 1996). Similar research has spurred an entire area of research examining intelligence from a neuropsychological perspective, which ties together many critical cognitive processes to behavior and areas of the brain.

Neuropsychological measures. The key to the neuropsychological approach is its basis in biological and psychophysiological theories, providing it with a strong set of theoretically driven tenets and, hence, making it a critical bridge between theories of intelligence that rely on the statistical approaches to categorize intelligence and those that rely on theoretical tenets to offer explanatory mechanisms of intelligence (Becker, et al., 2015). The neuropsychological approach offers deeper insight into the cognitive processes that underlie intelligence, and, as a result, offers new opportunities for the measurement of intelligence. It is important to recognize that while neuropsychology has an extensive history in clinical psychology, it remains a developing topic in the intelligence literature (Miller & Maricle, 2012). Therefore, it is important to consider how various approaches to intelligence from a neuropsychological perspective can be utilized together to better infer the structure of intelligence, its conceptualization, and how to best
operationalize it. This is a particularly important point, as a crucial critique of intelligence measures has been their lack of a theoretical foundation (Kaufman, 2000). Furthermore, it fits with the Parieto-Frontal Intelligence Theory (P-Fit), a popular neurocognitive theory of intelligence, which suggests that intelligence and its resulting behaviors can be arrived at through several different neurocognitive pathways (Haier & Jung, 2007; Jung & Haier, 2007). As such, it is important to consider how intelligence can manifest through various regions of the brain and therefore may not be limited to a single neuropsychological theory. For that reason, I review several neuropsychological research streams related to intelligence, mainly the planning, attention-arousal, simultaneous, and successive (PASS) theory of intelligence and research on the executive functions located in the prefrontal cortex.

PASS. One of the most prominent examples of neuropsychological research on intelligence is the PASS theory, which is applied frequently in developmental and educational psychology (Das & Varnhagen, 1986; Naglieri & Das, 1997, 2005; Naglieri et al., 2005). Based on prior neuropsychological research (Luria, 1966a, 1966b, 1973), the theory’s main tenet is that intelligence is comprised of four separate cognitive processes that interact with one another. Of these cognitive processes, there is planning; which entails processes around solving problems through strategic thinking and self-monitoring; attention; which involves maintaining focus on stimuli in the face of distractors; simultaneous processing, which involves organizing stimuli into coherent patterns and perceiving the relationships between stimuli; and successive processing; which includes arranging information into a sequential order (Naglieri et al., 2005). Much like executive attention, which has multiple forms of attention processes, the PASS theory identifies several modules of the brain necessary for problem solving. Each of these modules is associated with a separate area of the brain, including areas beyond just the frontal lobe (Das et
Furthermore, the PASS theory emphasizes that these four processes must operate within a person’s knowledge base, including both temporary and long-term knowledge, and ultimately influence a person’s knowledge acquisition (Naglieri & Das, 1997).

To assess the distinct cognitive processes of the PASS theory (i.e., planning, attention, simultaneous, and successive processing), Naglieri and Das (Naglieri, 2005; Naglieri & Das, 1997) developed the Cognitive Assessment System (CAS). For each of the processes, there are multiple subtests that vary in content (Das, 2002). The measures of simultaneous and successive processing incorporate information-processing tasks, whereas the measures of planning and attention include tasks associated with higher-order control process (Fein & Day, 2004).

The measures assessing planning in the CAS require the test taker to create and apply a plan of action and subsequently monitor the effectiveness of that plan during a task. As such, the measure plays a critical role in assessing a person’s problem-solving abilities. There are multiple measures of planning, including the matching numbers, planned codes, and planned connections tasks. As an example of one of these measures, the matching numbers task involves the test taker underlining any identical numbers in a series of numbers. As the task progresses, the number of digits in the series of numbers grows, increasing the complexity of the task. A unique aspect of planning measures is that they require an assessment of the actual strategies used to solve the task and are partially coded by an observer.

The attention measures are intended to assess a person’s ability to detect a particular stimulus while inhibiting distractors. In many ways, these measures are similar to the measures of attention included in executive attention batteries. An example measure of an attention measure is the receptive attention task, which requires the test taker to underline any pairs of letters that are the same based on a set of rules. For one part of the test, the test taker must
underline all letters that are the same and the same size (e.g., AA or aa), whereas in the second part the test taker must underline those letters that are the same (e.g., aA or Bb). Another measure is the expressive attention task, which is a variation of the Stroop test. For this task, the test taker is shown a visual image (e.g., the word blue) but must verbalize the correct response (e.g., the color of the text the word is written in).

The measures for simultaneous processing require the test taker to organize disparate pieces of information and arrive at valid conclusions from this organization. The operationalization of this system is similar to many of the operationalizations of fluid reasoning. For example, one measure is known as nonverbal matrices and follows the principles of the Raven’s Progressive Matrices. Similarly, another test known as the verbal-spatial relations task is similar to one of the Cattell Culture Fair Intelligence Tests and involves the test taker selecting a specific figure that conforms to a set of rules (e.g., select the figure with a triangle to the left of a circle).

The final system of PASS, successive processing, involves comprehending meaning from the order in which information is presented. Often, this requires a person to complete information in a specific order. For example, the sentence questions task entails the test taker reading a nonsense sentence (e.g., the blue yellows the orange) and then must answer a question about the sentence (e.g., who yellowed the orange?). Other tasks associated with successive processing require the test taker to repeat back words in various orders. The sentence repetition task involves the test taker repeating back sentences that have syntax but reduced meaning.

Taken together, these CAS tests rely on novel tasks and limit verbal and quantitative knowledge in order to get at a person’s high-level processes detached from prior knowledge (Naglieri & Bornstein, 2003). While the CAS was originally developed for children and
adolescents, it has been applied to older participants to predict performance on a complex performance task (Fein & Day, 2004). As such, the strong theoretical foundation and multidimensional approach of this measure could be a beneficial tool for research in the workplace going forward.

Executive functions. While the PASS theory considers several specific cognitive functions throughout the brain, there has also been considerable work examining various executive functions in the dorsolateral prefrontal cognitive ability (D-PFCA), an area of the brain frequently connected to intelligence (e.g., Jung & Haier, 2007). Although there have been various models of executive functions, generally they are conceptualized as cognitive processes associated with goal-directed behavior and controlled cognition (Banich, 2009; Fuster, 1997). To date, executive functions have been considered to cover a wide range of processes, including shifting mental sets, monitoring progress towards a goal, updating task demands, cognitive flexibility, and even working memory. While historically the work on executive functions has been piecemeal, examining many of these processes individually or in subsets, there is growing work to better understand how they correspond to the D-PFCA and why they operate in the same region of the brain (Alvarez & Emory, 2006; C. Blair, Zelazo, & Greenberg, 2005). This has resulted in a debate regarding whether executive functions are one unified system or if they are a loosely connected set of processes (e.g., Blair et al., 2005).

The lack of clarity in how the executive functions relate to one another can be seen in how they are measured as well. Numerous measures have been developed with the intention of assessing executive functions, but they are often developed and applied independently. However, more recently batteries of these prefrontal cognitive tasks have been utilized in several studies and found to predict performance in both work and academic settings (Higgins et al., 2007;
Evidence shows that these tasks are associated with various executive functions, such as manipulating and monitoring information, and encoding and retrieval of information in memory (for a review, see Higgins et al., 2007). One battery of D-PFCA developed by Higgins and colleagues (2007) aims to assess three categories of tasks: conditional associative tasks, working memory, and word fluency.

Conditional associative tasks assess a person’s ability to learn conditional behavioral rules required to complete a task. There are three measures of conditional associate tasks in the Higgins and colleagues’ battery, including a go/no-go task. In this task, the test taker is required to respond to certain prompts with either a go or no-go response depending on predetermined criteria. In this type of task, performance is predicated on the successful learning of the criteria and subsequently responding correctly. Other conditional associative tasks in the battery include a spatial and a non-spatial conditional associative task. For the spatial conditional associative task, the test taker is shown a set of five squares and five circles. The test taker is told that each square is associated with just one circle. In each trial, one circle is highlighted, and the test taker must guess which square corresponds with it. The test taker can guess until he/she correctly identifies the corresponding square. Trials continue until the test taker correctly identifies the cue word in ten consecutive trials or after completing 100 trials altogether. The non-spatial conditional associative task follows the same design except the test taker is shown five target words and one of five cue words. The test taker must then guess the one target word associated with the cue word before moving onto the next trial. Again, this continues until the test taker correctly identifies the cue word in ten consecutive trials or after completing 100 trials.

The measures used for assessing working memory from an executive function perspective focus on the test taker’s ability to store, update, maintain, and inhibit information
throughout a task. Within Higgins and colleagues’ D-PFCA battery, there are three working memory measures. One measure is the self-ordered pointing task. In this task, the test taker is shown a set of stimuli, which can be words, images, or figures. Typically, 12 stimuli are shown at once and the test taker must click on one stimulus, at which point all of the stimuli will be randomly reordered. The test taker must then click on another stimulus without clicking on a previously selected stimulus. The test taker is able to click as many times as there are stimuli. Another measure in the battery is known as the randomization task. For this, the test taker is asked to randomize the order of letters in a given span. For example, the test taker might be asked to randomize the letters from J to M. Once the test taker correctly performs that task, he/she is asked to randomize a new span of letters, one letter longer than the previous span. If the test taker does not use all the letters, uses the wrong letters, or puts the letter in its original sequence (e.g. OP), the trial is considered incorrect. The test taker continues until they get two trials in a row wrong or complete a 14-letter span. The final working memory task is the recency discrimination task. For this task, the test taker is briefly shown a sequence of six or eight words. After the sequence is displayed, two of the words are displayed again and the test taker must identify which of the two words was shown most recently.

Finally, word fluency tasks require the test taker to produce as many words he/she can in a set time based on a rule or set of rules. For the D-PFCA battery by Higgins and colleagues, there is only one task of word fluency. For this task, the test taker is asked to generate as many words that start with a specific letter or set of letters, such as “ST”. The test taker must generate as many words as he/she can within five minutes and cannot use inflections of a submitted word, such as plurals or past tense.
A critical feature of Higgins and colleagues’ measure of executive functions is the fact that it is a battery of multiple cognitive processes. While each of these measures assesses an independent cognitive process, scores are only reported for a composite D-PFCA score. Therefore, while the measure uses discrete tasks to evaluate the cognitive processes, its score reflects a broad construct of information processing. This is an important distinction since the measure contains a diverse array of tasks and stimuli to assess one overall construct. The diverse tasks found in this measure offers a representative set of possible tasks that can be used in research, although other tasks exist and have been used to measure executive function (e.g., Culbertson, Huffcutt, & Goebl, 2013; Stermac-Stein, 2014). Much like Higgins and colleagues’ D-PFCA battery, the tasks in these other measures seek to assess high-level cognitive processes that are used for controlling processes for goal attainment, and therefore share many similarities to the tasks used in the D-PFCA battery.

Neuropsychology measures and performance. Most research involving neuropsychological measures of information processing have been used in diagnostic settings, intended for the use of interventions in children or clinical populations (e.g., Das, 2002; Naglieri, 1999; Peterson, Finn, & Pihl, 1992; Peterson, Pihl, Higgins, Séguin, & Tremblay, 2003; Séguin, Nagin, Assaad, & Tremblay, 2004). However, a handful of studies with adults at work and in educational settings have provided promising findings about how these measures predict performance. For example, Higgins and colleagues (2007) found that the D-PFCA battery is related to job performance in both managerial and factory floor jobs. They also found that the battery was correlated with academic performance, which was replicated by Sabet and colleagues (2013). Research on another neuropsychological measure, the CAS, has also offered evidence that these measures can predict a person’s ability to acquire a complex skill (Fein &
Day, 2004). Taken together, the research on neuropsychological measures demonstrates the positive relationship between these measures and performance.

**Neuropsychology measures and group score differences.** Research on group score differences between African Americans and Caucasians is extremely limited for neuropsychology measures, with only one study currently published on this topic in the I-O psychology literature. Sabet and colleagues (2013) found that the score differences between African Americans and Caucasians was a $d$ of .15, which is considered a small difference. While these results offer an encouraging start, extensive research is needed to better understand group score differences on these measures in I-O psychology.

**Summary of neuropsychology measures.** While there are other neuropsychological measures that are currently being applied in clinical settings or being developed today, many of them have yet to be used to predict performance. Still, taking together the current research, the neuropsychological approach offers a unique and theoretically sound approach to understanding and measuring intelligence. By grounding cognitive processes in regions of the brain, it is possible to better understand how these processes relate to each other and how information processing measures contribute to the overall construct space. Specifically, by using neuroimaging, it may be possible to understand the unique contributions of each measure in assessing information processing (Haier & Jung, 2007; Jung & Haier, 2007).

**Summary of the cognitive science approach.** Measures from the cognitive science approach have been used sparingly in the I-O psychology literature, however, researchers have commented on its potential for use in the field (e.g., Agnello et al., 2015; Scherbaum et al., 2012). With the handful of studies demonstrating the potential for these measures to successfully predict performance in several settings, it is important to continue research on these measures.
The strong theoretical foundation and multidimensional approach of these measures could be a beneficial tool for research in the workplace going forward.

**Hypotheses**

Research on information processing measures has expanded in recent decades with renewed interest in the psychometric approach (e.g., Schneider & McGrew, 2012; Schneider & Newman, 2015) and new advancements in cognitive sciences (e.g., Ardila, 1999; Becker et al., 2015). This has led to a strong foundation of research examining information processing measures and many of their key features, such as their predictive power and how people from different groups perform on them. However, recent calls for additional research on these measures (e.g., Fagan & Ployhart, 2015; Reeve et al., 2015) demonstrates that there are still deficits in the collective knowledge about information processing measures. As such, the goal of the current meta-analysis is to aggregate extant findings on these measures to spur discussion about their utility and encourage further research. However, the fact that research on these measures is still evolving, especially in the cognitive sciences, means that testing some conditional effects in a meta-analysis is currently impractical due to low sample sizes. For that reason, I hypothesize several main effects and moderators about information processing measures but do not propose moderation effects for all main effects. While this limits the tests that can be conducted, it demonstrates the importance of the current meta-analysis as a preliminary review of the existing findings in the field and highlights specific areas requiring more research going forward.

**Relationship Between Measures of Information Processing and Performance**

Based on the extant literature on information processing, there is a clear basis to expect information processing measures to correlate with performance in multiple domains, such as job
performance, job-oriented training performance, and academic performance. Foremost, information processing reflects a person’s ability to attend to critical information in the environment, reflect on and manipulate information, deduct rationale conclusions from information, and apply those conclusions to additional stimuli perceived in the environment. These core cognitive functions are critical for numerous behaviors required for success across numerous domains, especially those that require complex behaviors or those that require high attentional resources, such as those that involve processing information in large amounts or over an extended time (Kanfer & Ackerman, 2005). Therefore, it is expected that information processing predicts performance on these types of tasks. Past research supports this notion, with evidence demonstrating information processing predicts the ability to multi-task (e.g., Bühner et al., 2006), performance on a flight simulator (Sommer, Häusler, Koning, & Arendasy, 2006; Van Benthem & Herdman, 2016), and performance on a managerial simulation (Bosco et al., 2015).

Furthermore, information processing is considered to represent the core set of cognitive processes associated with basic intelligent functioning (Fagan & Ployhart, 2015). For that reason, it is believed to be the fundamental process for learning skills, abilities, and knowledge (Ackerman, 1996). By its nature, information processing provides the basis for a person’s declarative and procedural knowledge, which are at the core of performance on various tasks at work (Kanfer & Ackerman, 2005) and in education (Kuncel, Hezlett, & Ones, 2004). As such, information processing also has an indirect effect on performance through one’s ability to learn new information (Ackerman & Kanfer, 1993). This is demonstrated in past research that shows information processing predicts knowledge and skill acquisition (e.g., Moran, 1986; Roznowski et al., 2001) and subsequent transfer to performance outcomes (e.g., Bergerson & Gustafsson, 2011; Fein & Day, 2004). This aligns with theory that suggests knowledge acquisition is the
main causal factor predicting job performance across most jobs, including those with low complexity (Schmidt & Hunter, 1992). The combination of information processing predicting learning and performance on complex tasks leads to the proposition that information processing should also predict overall performance at both work and school, which has been supported in a multitude of prior studies as described in the previous sections (e.g., Henderson, 1979; Higgins et al., 2007; Rohde & Thompson, 2007; Sabet et al., 2013). Based on these previous findings, it is hypothesized that:

_Hypothesis 1:_ Information processing measures will be positively correlated with (a) job performance, (b) job-oriented training performance, and (c) academic performance.

**Group Score Differences on Measures of Information Processing**

Research on group score differences is a critical area for applied psychology fields, particularly I-O psychology. Differences on intelligence scores can ultimately lead to disparities in outcomes, such as admission into college and selection for a job, that can have substantial impact on a person’s life. For that reason, group score differences and its related concept of adverse impact are important legal issues, especially in the United States. Furthermore, organizations often seek to improve diversity in their workforce for a wide range of “business, social, or ethical reasons” (Ployhart & Holtz, 2008, p. 153). Most of the research in I-O psychology has focused on differences between African Americans and Caucasians as they have been systematically presented as the largest group difference (Hough et al., 2001; Neisser et al., 1996). For that reason, I focus on group score differences between African Americans and Caucasians in the current study. In addition, because each country has its own social, cultural, and legal histories that influence how group score differences manifest and are examined (Outtz & Newman, 2010), I only examine group score differences within the United States.
Group score differences on measures of intelligence are typically represented as a standardized mean score difference with Cohen’s $d$. Though specific group score differences vary by measure and may even further vary by setting (Outtz & Newman, 2010; Roth et al., 2001), several prior meta-analyses and narrative reviews have referenced a $d$ of 1.00 (or one standard deviation from the mean) between African Americans and Caucasians on measures of intelligence (Hough et al., 2001; Hunter & Hunter, 1984; Ployhart & Holtz, 2008; Roth et al., 2001; Sackett & Wilk, 1994; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997). This value has been demonstrated across a wide range of studies, samples, and measures, thus becoming a standard benchmark to compare group differences for any given intelligence measure against. Most notably, Roth and colleagues (2001) conducted a meta-analysis of intelligence measures across several different samples. They found that intelligence measures across various samples had a $d$ value of 1.10 ($k = 105$) with a lower limit to its 95% confidence interval of 1.06, suggesting that the group score differences on these measures were most likely above a 1.00. In addition, they found that in an education sample, the $d$ value was 1.12 ($k = 148$) with a lower confidence limit of 1.09, and in an industrial sample they found a $d$ value of .99 ($k = .88$) with a lower limit of .88. Finally, they examined a specific sample of applicants that had taken the Armed Services Vocational Aptitude Battery, a measure believed to mainly assess crystallized intelligence (Roberts et al., 2000), and found a $d$ value of 1.19 ($N = 212,238$). An analysis across three studies of the Wonderlic, another measure mainly assessing crystallized knowledge, found a $d$ value of 1.09 and a lower confidence limit of .89 (Bosco et al., 2015). Taken together, these findings suggest that while there is variability in group score differences, they are consistently high with the reported $d$ value of 1.00 an appropriate benchmark across intelligence measures.
The source of these group score differences remains a fiercely debated topic (Neisser et al., 1996). While research continues to investigate the determinants of these differences (e.g., Cottrell et al., 2015; Moore, 1986), researchers argue that score differences between groups can be a result of contamination and/or deficiencies in measures of knowledge (Fagan, 2000; Fagan & Holland, 2002; Goldstein et al., 2002). Knowledge-based measures, by their nature, rely on information that may differ based on a person’s schooling and cultural background (Fagan & Holland, 2002) and even differences in his/her family’s income, maternal education, and maternal verbal ability and knowledge (Cottrell et al., 2015; Outtz & Newman, 2010). When this is coupled with findings that show it is possible to reduce group mean differences by removing item content that involves knowledge that varies by race (Freedle & Kostin, 1997; Fagan & Holland, 2007; Malda, et al., 2010), there is clear support that group score differences are at least partially a result of items that require prior knowledge.

Researchers suggest that when basic cognitive functions are assessed, like those in information processing measures, a person’s raw intellectual capability is being measured instead of the outcomes of that capability (i.e., knowledge) (e.g., Fagan, 2000; Fagan & Ployhart, 2015; Sternberg, 1997, 2000). At their core, information processing measures are intended to assess a person’s ability to solve novel problems and therefore should minimize prior knowledge in the assessment process. Supporting this notion, Hough and colleagues (2001) found that measures not relying on prior knowledge generally had lower mean differences between African Americans and Caucasians than knowledge-based measures. Furthermore, findings from both the psychometric (e.g., Colom & García-López, 2002; Klein et al., 2007; Nenty & Dinero, 1981) and cognitive sciences approach (e.g., Bosco et al., 2015; Naglieri, 2005; Sabet et al., 2013) have consistently found information processing measures have smaller group difference scores than
traditional measures of general intelligence or intelligence-as-knowledge. For that reason, the following is hypothesized:

*Hypothesis 2*: Information processing measures will have smaller group scores differences between African American and Caucasian test takers than traditionally found on intelligence measures (i.e., $d = 1.00$).

**Moderators**

In addition to main hypotheses regarding information processing, it is feasible that there will be additional variance to be explained by conditional effects. While many different variables may influence the relationship between information processing and performance and group score differences on these types of tests, this meta-analysis will focus on three characteristics of the test as potential moderators: the theoretical approach used to develop the test, the diversity of task types within a test, and the level of language knowledge required for the test.

**Theoretical approach.** The psychometric and cognitive science approaches to information processing share many commonalities, to the point that researchers have tried to map them onto each other (e.g., Deary, 2005; Engle, 2018; Kyllonen & Christal, 1990; Ren, Schweizer, Wang, Chu, & Gong, 2017; Schneider & McGrew, 2012). Researchers have even begun to integrate cognitive science research into psychometric models, with the Cattell-Horn-Carroll model including factors corresponding to working memory and controlled executive attention (McGrew, 2005). Similarly, recent advancements have led test developers to incorporate measurement strategies from both approaches, such as assessing the cognitive processes that underlie psychometric measures and applying advanced factor analysis methods on neuropsychological measures (e.g., Fiorello, Hale, Snyder, Forrest, & Teodori, 2008; Flanagan, Alfonso, Ortiz, & Dynda, 2010; Primi, 2002). Though it is obvious that there are
shared features between the functions examined in cognitive sciences and constructs such as fluid reasoning, the exact relationships between these constructs remain unclear (Kent, 2017). Furthermore, there are several critical distinctions between the theoretical approaches that make them unique in both their conceptualization and operationalization of information processing. Specifically, the two approaches mainly differ in their theoretical grounding, coverage of the construct space, and item designs.

Perhaps the most critical difference between the theoretical approaches is their reliance on a priori theory in conceptual and operational development. Whereas the psychometric approach relies on results from statistical analyses to identify the structure of information processing, the cognitive sciences approach utilizes multiple pieces of data founded in the biological and neurophysiological sciences, such as neuroimaging and traumatic brain injury studies (Sabet et al., 2013). This means that the psychometric approach inherently begins its evaluation of information processing and its components at a higher level of abstraction than the cognitive sciences. It also indicates that the psychometric approach has less integration of theories from other disciplines at its foundation since factor analysis provides the initial insights for understanding information processing compared to the cognitive sciences, which incorporate constructs and methodologies from multiple disciplines throughout theory development. To be clear, there are examples in the psychometric approach that are founded in strong theory, with Hebb’s original work on different biological underpinnings of intelligence (Brown, 2016) and Cattell’s attempts to develop a culture-free intelligence measure to assess reasoning ability rather than knowledge (Cattell, 1984) at the crux of the psychometric approach to information processing. However, much of the theory development from the psychometric approach has relied on leveraging statistical analyses to better differentiate intelligence constructs, such as how
fluid reasoning differs from general intelligence (e.g., Gustafsson & Åberg-Bengtsson, 2010; Jensen, 1998), rather than integrating other theories to evolve the constructs, although this has begun to change in the last decade (Flanagan et al., 2010; Primi, 2014).

The differing degree of a priori theory integration between these two approaches also carries over to their development of information processing measures and the range of the construct space they assess. While the psychometric approach is founded in theories related to information processing, the origins of its constructs, specifically fluid reasoning, is not originally based on physiological research of the brain. Alternatively, the cognitive science approach focuses on developing constructs and subsequent measures that are based on physiological and neuropsychological research. This creates a difference in the level of abstraction at which the measures begin. While the psychometric approach is more statistically abstract, the cognitive science approach is more proximal to the intended cognitive processes of the brain. This has led researchers to argue that measures from the cognitive sciences are able to capture a broader range of information processing functions compared to other approaches because they are more directly linked to a variety of cognitive processes (Conway et al., 2003; Higgins et al., 2007). In doing so, the cognitive science approach should reduce potential construct underrepresentation and have stronger relationships with performance outcomes (Daniel, 1997; Messick, 1995).

Finally, measures from the cognitive science approach and psychometric approach typically differ in the design of their items. Specifically, measures from the psychometric approach often include static stimuli, such as the matrix items typically used to measure fluid reasoning. Although the items on these measures require reasoning ability, the test taker only needs to identify a right response from multiple choices, much like an achievement test, for their ability to be assessed. This makes it difficult to know the explicit processes underlying a test.
taker’s response since the operation of the cognitive processes is not observable. On the other hand, measures from the cognitive science approach rely on dynamic behavioral based responses (Kyllonen, 2002; Miller & Maricle, 2012). These measures include maintaining a string of numbers in memory while identifying whether a series of math equations are true or false (i.e., working memory tasks; Unsworth, Heitz, Schrock, & Engle, 2005), clicking when certain stimuli are presented on the screen (i.e., conditional associative learning tasks) (Peterson et al., 2003), or developing strategies for matching the same number in a row of numbers (i.e., planning tasks) (Naglieri & Das, 1997). As a result, measures from the cognitive science approach offer more direct observations of the cognitive processes in question and have a high degree of correspondence between the construct of interest and the unit of measurement, an important factor for improving construct validity (e.g., Guion, 2002; Stone-Romero, 1994). Specifically, with an increased ability to assess the psychological construct directly, a test acts more as a direct “sample” of the target behavior rather than an indirect “sign” (Cronbach & Meehl, 1955). This improves the accuracy of the test in assessing the underlying construct and can help eliminate unwanted variance in the measurement process, potentially improving the predictive validity of the measure.

Taken together, the cognitive science approach should be more theoretically bound, cover a wider scope of the information processing construct space, and include test items that more closely reflect the actual operation of the target cognitive processes compared to the psychometric approach. These features of the cognitive science approach should subsequently reduce contamination and deficiencies within the testing process by capturing specific cognitive functions through high fidelity measures. Therefore, it is anticipated that while psychometric measures of information processing do predict performance, they will not be as successful as
measures from the cognitive sciences. Supporting these notions, research has shown that neuropsychological measures of information processing can predict performance above and beyond fluid reasoning (Higgins et al., 2007). Based on these findings, it is predicted that:

*Hypothesis 3:* Cognitive science measures of information processing will have a stronger relationship with (a) job performance, (b) job-oriented training performance, and (c) academic performance than psychometric measures of information processing.

Unfortunately, theories of adverse impact and group score differences are still only emerging (e.g., Cottrell et al., 2015) and there is little conclusive research on group score differences, particularly for measures from the psychometric and cognitive science approaches. However, there is research showing some measures from the psychometric approach, including the widely used Raven’s Progressive Matrices, still exhibit group score differences (e.g., Colom & Garcia-Lopez, 2002). Early research has also shown that a measure of executive attention had lower score differences than the Progressive Matrices, though the difference was not significant (Bosco et al., 2015). For that reason, it is important to quantitatively evaluate how these theoretical approaches may differentially influence group scores.

The fact that the cognitive science approach is founded soundly in physiological research may have implications for how the two approaches manifest group score differences. Although the causes of groups score differences are still not completely understood, there is growing research that suggests genetic and biological factors do not fully explain differences in group scores (for reviews see Neisser et al., 1996 and Nisbett et al., 2012). Because the theory and measurement of the cognitive science approach originates more closely to the physiological domain than the psychometric approach does, it is possible that the cognitive science approach
can reduce measurement deficiencies and contamination that are associated with irrelevant error contributing to group score differences. For that reason, I hypothesize the following:

*Hypothesis 4:* Cognitive science measures of information processing will have smaller group score differences between African American and Caucasian test takers than psychometric measures of information processing.

**Diversity of task types.** Attempts to optimize construct validity of information processing measures are ultimately aimed at increasing the alignment between the conceptualization and operationalization of a construct. As noted above, this could involve using measures that directly reflect the behaviors of information processing and, thus, sample the domain of interest. In the intelligence literature, researchers have also noted the importance of reducing the potential error caused by only sampling intelligence with a limited number of tasks, known as the psychometric sampling error (Jensen & Weng, 1994). From this perspective, the use of diverse tasks in a measure of intelligence helps to remove unwanted error by aggregating across multiple tasks, which ultimately helps better triangulate the focal construct (Reeve & Blacksmith, 2009). Similarly, others have argued that construct validity of intelligence measures is limited when using a narrow operationalization of intelligence (Kretzschmar, Neubert, Wüstenberg, & Greiff, 2016). Research has supported this notion by demonstrating that measures of fluid reasoning (Beauducel, Liepmann, Felfe, & Nettelnstroth, 2007) and critical problem solving (Kretzschmar et al., 2016) that use a diverse set of tasks to measure the focal construct have stronger correlations with hypothesized variables of interest than measures with a narrow set of tasks. By increasing the construct validity of the test through a diverse set of tasks should also lead to increased predictive validity since potential error is being removed from the measure.
While a diverse set of tasks has the potential to be beneficial for the construct validity of information processing measures, the rationale may differ slightly for measures from the psychometric and cognitive science approaches. Specifically, the diversity of tasks used in psychometric measures often differ based on the format of the item (e.g., non-verbal matrix, verbal spatial relations) whereas the diversity of tasks used in cognitive science measures often reflects multiple cognitive subsystems being measured. As an example, the Cattell Culture Fair Test is a psychometric measure of information processing that consists of four subtests (Cattell, 1973). Each subtest is intended as a measure of fluid reasoning, though each is a different type of perceptual task in order to avoid reliance on a single ability. As such, the diversity of tasks associated with this test and others from the psychometric approach are specifically aimed at reducing the psychometric sampling error of the measure. On the other hand, a test battery of D-PFCA from the cognitive science approach includes multiple measures to assess distinct cognitive processes that are all associated with the D-PFCA (Higgins et al., 2007). Therefore, while there are broad and diverse tasks, much like the psychometric approach, these diverse tasks attempt to sample a broader scope of cognitive processes included in the D-PFCA. Similarly, a measure of executive attention includes multiple tasks associated with assessing the various cognitive processes associated with both complex and simple attention tasks (e.g., Bosco et al., 2015; Hutchinson, 2007). As such, the cognitive approach should still reduce psychometric sampling error but do so as a result of minimizing measurement deficiencies of the focal construct.

Although the outcome of having diverse tasks in the measure is expected to improve construct validity and, as a result, predictive validity for both the psychometric and cognitive science approaches, the underlying rationale diverges. To avoid conflating the effect of diverse
tasks across the two theoretical approaches, I offer separate hypotheses for each approach. As such, I predict the following:

**Hypothesis 5:** Psychometric measures of information processing with a diverse set of item types will have a stronger relationship with (a) job performance, (b) job-oriented training performance, and (c) academic performance than psychometric measures of information processing that do not have a diverse set of item types.

**Hypothesis 6:** Cognitive science measures of information processing with a diverse set of item types will have a stronger relationship with (a) job performance, (b) job-oriented training performance, and (c) academic performance than cognitive science measures of information processing that do not have a diverse set of item types.

**Language knowledge requirements.** Researchers have begun to promote the use of information processing measures for their reduced reliance on prior knowledge (e.g., Agnello et al., 2015; Scherbaum et al., 2012). In particular, for testing contexts in which knowledge is not required, these measures can potentially reduce contamination from irrelevant testing content (Fagan, 2000; Fagan & Holland, 2002; Goldstein et al., 2009). However, even within information processing measures there is variation in the degree prior knowledge is needed to respond to the items. While all information processing measures are intended to assess the cognitive processes that underlie problem solving and reasoning, the format of the items differ across tests. Graphical stimuli, such as those used in the Raven’s, are commonly associated with information processing measures, but other measures utilize vocabulary as the stimuli. Examples of measures with vocabulary stimuli come from both the psychometric and cognitive science approaches. For instance, the Baddeley Reasoning Test is a psychometric test that requires test takers to determine whether a series of sentences are true or false based on a set of prespecified
rules as quickly as possible (Baddeley, 1968). Similarly, the word fluency test, which is a measure from the cognitive sciences, requires test takers to produce words that meet a set of specific guidelines, such as the words must start with the letters ST (Higgins et al., 2007). In both of these examples, the test taker must have knowledge of the language the test is being administered in.

While all tests have some language requirements to the extent they include instructions, tests that have language infused into their items require the test taker to have at least some language knowledge to successful respond to individual items. Most measures of intelligence-as-knowledge require language knowledge, which can be advantageous in some circumstances. For example, measures of verbal ability can be useful in hiring scenarios in which an organization is filling a position requiring language comprehension (e.g., Wee, Newman, & Joseph, 2014). Still, research has demonstrated that variation in test taker knowledge of the specific vocabulary on a test can lead to differences in performance on a measure of intelligence (e.g., Fagan & Holland, 2002, 2007, 2009; Freedle & Kostin, 1997). For this reason, researchers have suggested that intelligence measures reduce the language demands associated with intelligence tests and instead use graphical stimuli (Naglieri, 2005), particularly in situations where language knowledge is not required for successful performance on the criterion. In cases where language knowledge is important for the job, it is important for test developers to ensure the level of language aligns with the requirements of the job. This corresponds with appeals to utilize more non-entrenched stimuli, such as novel or abnormal stimuli, in testing situations where language is not critical to ensure all test takers have similar experience with the test content (e.g., Bokhorst, 1989; Sternberg, 1982; Tetewsky & Sternberg, 1986). In the case of information processing measures,
doing so can eliminate irrelevant content from the measure and reduce potential contamination of the measure. Therefore, I propose the following:

**Hypothesis 7:** Information processing measures that require less language knowledge will have a stronger relationship with (a) job performance, (b) job-oriented training performance, and (c) academic performance than information processing measures that require more language knowledge.

While there are information processing measures from both theoretical approaches, they are overall limited in number. Therefore, while it is important to investigate the effect of language knowledge on performance, it is currently not feasible to assess its effect within the various theoretical approaches. Similarly, the current dearth of research makes it impractical to test for the effect of task diversity and language knowledge on group score differences.

**Method**

To examine these hypotheses, I applied meta-analytic methods to aggregate the extant data on information processing measures. Meta-analysis offers a unique quantitative approach for combining and reporting findings across multiple samples, which allows for a larger sample size and higher statistical power (Rosenthal & DiMatteo, 2001). Specifically, the effect sizes from individual samples are combined and corrected based on the various errors associated with the individual samples, such as measurement error in the criterion, offering more accurate and credible conclusions than a single study can provide (Hunter & Schmidt, 2004). In addition, any variance not explained by the sampling variance can be further evaluated by examining the potential effect of moderators (Borenstein, Hedges, Higgins, & Rothstein, 2009). To do so, each effect size within a meta-analysis is coded to determine the relevant statistical data and sample characteristics that can be aggregated and compared to make big picture inferences about the
existing data. It is important to note that meta-analysis findings are not meant to be conclusive assessments of a research area, regardless of the scope of the meta-analysis (Aguinis, Pierce, Bosco, Dalton, & Dalton, 2011). Instead, they are an important stage in the evolution of a research area meant to summarize current knowledge, evaluate the standing of the field, and spark new research questions.

For the current meta-analysis, I utilized Hunter and Schmidt’s (2004) methodology for estimating sampling variances, correcting for measurement error in the criterion, and weighing effect sizes based on their sampling variance. To test my main effect and moderation hypotheses, I used a multi-level analysis in order to address potential issues of non-independent effect sizes (Bowman, 2012; Hox, 2010). The following sections offer an overview of the steps I conducted for the meta-analysis, including how I operationalized the focal variables, the process for conducting the relevant literature search, the inclusion and exclusion criteria for the meta-analysis, and the process for meta-analyzing the effect sizes.

**Operationalization of Key Variables**

**Information processing.** Information processing measures cover a wide spectrum. As mentioned above, both the psychometric and cognitive science approaches to intelligence offer similar yet distinct methods for assessing information processing. At their core, these measures are focused on assessing a person’s ability to perceive, integrate, and reason with information to make decisions and rely on stimuli that minimize prior knowledge to reduce contamination of other constructs. As such, this meta-analysis will include measures that most closely reflect the information processing domain, as conceptualized in this paper. A scale was determined to be a measure of information processing based on 1) the construct validity of the measure and 2) the intended construct of the scale by the test developer. Table 1 offers an overview of common
psychometric information processing measures and Table 2 offers an overview of common
cognitive science information processing measures, which were included in this meta-analysis.
Furthermore, while this meta-analysis includes information processing measures that are subtests
or dimensions of larger measures when they are reported as independent scales, it does not
include studies that aggregate information processing measures with knowledge measures. This
approach limits the number of effect sizes included in the meta-analysis, and potentially reduces
the statistical power of the findings, but allows for the most accurate test of the hypotheses as it
will only examine those measures that most closely reflect the information processing domain, as
conceptualized in this paper.

**Job performance.** In defining job performance, researchers often highlight that it is the
behaviors one engages in that are related to the goals of the organization (e.g., Campbell,
McHenry, & Wise, 1990; Motowidlo, Borman, & Schmit, 1997). This definition shifts attention
from solely the results of employees’ actions to the actions themselves, and their value to the
organization (Motowidlo, 2003). A large range of behaviors fall under the scope of this
definition, leaving much flexibility in how job performance is measured. Previously, job
performance was treated mainly as a single, overall factor. However, more recently researchers
have argued that one factor does not offer enough conceptual explanation of performance. There
have been various models that have depicted job performance as multidimensional (e.g.,
Campbell, McCloy, Oppler, & Sager, 1993; Viswesvaran & Ones, 2000). In the I-O psychology
literature (Hough & Oswald, 2000; Sackett & Lievens, 2008) and intelligence research
(Gonzalez-Mulé et al., 2014), performance is frequently separated into three major dimensions
referred to as task performance, organizational citizenship behavior (OCB), and
counterproductive workplace behavior (CWB) (e.g., Dalal, 2005; Lievens, Conway, & De Corte, 2008; Robinson & Bennett, 1995; Rotundo & Sackett, 2002; Viswesvaran & Ones, 2000).

Task performance is associated with those behaviors that directly impact the technical core of a job or organization (Borman & Motowidlo, 1993). Measures of task performance can be assessed using a wide array of methods, including supervisor and peer ratings of task performance as well as objective measures, such as measures of productivity, work samples, or simulations (e.g., Knapp et al., 2001). Objective measures of task performance, particularly work samples and simulations, are intended to reflect tasks completed on the job and can be administered to both employees and non-employees alike. For example, laboratory studies have previously shown that information processing predicts performance on management simulations (e.g., Bosco et al., 2015), multi-tasking simulations (Hambrick, Oswald, Darowski, Rench, & Brou, 2010), and flight simulators (e.g., Sommer et al., 2006), all of which were used to parallel task performance on the job. Therefore, studies that use task performance measures intended to simulate job performance were included in the meta-analysis, including those administered to non-employees.

Often contrasted with task performance are OCBs and CWBs. OCBs are considered behaviors that contribute to the success of the organization through their impact on the social and psychological environment, such as spreading goodwill and endorsing the organization’s goals (Rotundo & Sackett, 2002). Alternatively, CWBs are those behaviors that harm the goals of the organization or employees, such as sabotage, aggression, and even withdrawal (Bowling & Gruys, 2010). While intuitively these types of behaviors appear to be related to personality traits, such as conscientiousness and emotional stability, there is evidence that information processing predicts both OCBs (e.g., Côté & Miners, 2006) and CWBs (Dilchert, Ones, Davis, & Rostow,
One potential explanation for this link is that moral reasoning is founded on a person’s ability to reason and solve problems more generally (e.g., Derryberry, Wilson, Snyder, Norman, & Barger, 2005). Researchers hypothesize that a person’s likelihood to engage in behaviors such as OCBs and CWBs is influenced by the ability to consider the moral appropriateness, and subsequent rewards and consequences, of their actions (Dilchert et al., 2007). According to this perspective, the ability to anticipate the outcomes of a behavior can be indirectly affected by one’s information processing ability. Therefore, for the purposes of the current study, job performance was operationalized as overall job performance, task performance, OCBs, and CWBs.

**Job-oriented training performance.** At the core of job-oriented training effectiveness is changes in the trainee and, as a result, changes in the organization (e.g., Aguinis & Kraiger, 2009). When considering changes in the trainee, a key metric of success is trainee learning. Within the Kirkpatrick Model (Kirkpatrick, 1996), learning is focused on the knowledge gained and the changes in the trainee’s behavior as a result of the training. Similarly, Kraiger, Ford, and Salas (1993) proposed a tripartite model of learning, which includes cognitive learning, skill-based learning, and affective learning. In this model, cognitive learning references the acquisition of declarative knowledge whereas skill-based learning relates to procedural knowledge. In contrast, affective learning involves changes in opinion or attitudes, such as motivation or self-efficacy. While all of these categories of learning have important implications for the transfer of training to job-related behaviors, of interest for intelligence research are those training outcomes associated with the direct learning of knowledge and skills. For example, research has demonstrated that information processing predicts skill acquisition for apprentices during their first year of training (e.g., Moran, 1986) and knowledge-based learning following a
training session (e.g., Fein & Day, 2004). Therefore, I will focus on cognitive and skilled-based learning in the context of the current study.

**Academic performance.** Academic performance offers an important metric akin to job-related behaviors. While some have questioned the practical relationship between academic performance and real-world tasks (e.g., Sternberg & Wagner, 1993), others have highlighted the importance of declarative and procedural knowledge for academic success and how that knowledge closely parallels the knowledge and skills required for success in other life domains, including work (e.g., Kuncel et al., 2004). As such, it is unsurprising that academic performance has been found to predict future job performance across a wide spectrum of studies (Roth, BeVier, Switzer, & Schippmann, 1996). Academic performance can be assessed using grades from a single exam or assignment, a single class, or student’s cumulative grade point average. Although all these measures can be effective for assessing academic performance, researchers typically prefer aggregated measures of academic performance as they are more reliable indicators (Frisby, 2001; Robbins et al., 2004). Specifically, aggregated measures typically reflect performance over time, domains, and/or raters allowing for more consistency in the measurement and have been used in prior academic research (Higgins et al., 2007; Noftle & Robins, 2007). For the current study, most measures of academic performance reflected students’ aggregated grades over time within a class (e.g., overall course grade, end-of-year exams) or across classes (e.g., GPA), while measures of academic performance based on single exams or assignments were excluded. To date, numerous studies have demonstrated the relationship between information processing and academic performance, including final grade in a course (e.g., Krumm et al., 2012), overall GPA (e.g., Wüstenberg, Greiff, & Funke, 2012), and end-of-year exam scores (e.g., Furnham, Zhang, et al., 2006).
Literature Search

To identify studies relevant for the meta-analysis, a comprehensive search of available research databases was conducted. The process included a search of academic databases, including EBSCOhost, Google Scholar, JSTOR, PsycARTICLES, and PsycINFO. Various combinations of keywords were used as search terms, including the following: intelligence, cognitive ability, information processing, neuropsychology, fluid reasoning, executive attention, executive function, working memory, PASS, performance, academic performance, GPA, job performance, task performance, CWB, OCB, job-oriented training performance, race, group score differences, and adverse impact. Specific measures from the psychometric and cognitive science literatures (e.g., Cattell’s Culture Fair Test, CAS, Raven’s Progressive Matrices) were also used as keywords. Furthermore, articles of several key journals in the field (e.g., Journal of Applied Psychology, Educational and Psychological Measurement, Intelligence, Personnel Psychology, Journal of Applied Developmental Psychology, Cognition, Journal of Clinical Psychology) along with any relevant review articles, chapters, and books were reviewed to identify additional studies. Reverse searches were conducted from the reference section of prior meta-analytic studies of intelligence (e.g., Bertua et al., 2005; Gonzalez-Mulé et al., 2014; Hough et al., 2001; Hülsheger et al., 2007; Kuncel et al., 2004; Postlethwaite, 2012; Salgado et al., 2003; Ziegler et al., 2011) to isolate information processing measures used in those studies. Using this initial set of articles, chapters, and manuals, a reverse search of their reference sections was conducted to identify additional studies.

In addition to collecting published data, attempts were made to overcome the “file drawer problem,” which reflects the tendency for significant results to be published more often than non-significant results (Dickersin, 2005; Rosenthal, 1995). To do so, searches were also
conducted on ProQuest Dissertations and Theses with the same search terms as described above. Manual reviews of conference programs for key associations in the field (e.g., Academy of Management, American Psychological Association, Society of Industrial-Organizational Psychology) were conducted for presentations or posters on topics related to information processing measures. Technical reports and manuals were reviewed for measures or data (e.g., Aamodt, 2004). I also contacted major test publishers of cognitive ability tests and prominent researchers working on information processing measures to inquire about any unpublished samples they may have. These search strategies cumulated in the database of possible studies to be used in the meta-analysis.

**Inclusion and Exclusion Criteria**

After identifying studies from the literature search that could be potentially included in the meta-analysis, the studies were evaluated against a series of a priori criteria to determine if they would be retained for data analysis (see Figure 1 for inclusion and exclusion criteria). For a study to be retained, it had to include an information processing measure, as operationalized above. In addition, the study had to present data about the measure and 1) its relationship with job performance, job-oriented training performance, or academic performance, as operationalized above, and/or 2) provide data on group scores for African Americans and Caucasians. Next, the data presented in the study needed to include statistics that could be transformed into a common effect size. For studies examining the relationship between information processing and performance, the focal effect size was Pearson correlation coefficients because most intelligence research reports correlations. Therefore, a study had to include the correlation between the measures or other statistics that could be transformed into a correlation (e.g., *t*-values, *F*-ratios) to be included. When reporting on group scores for a
measure, there needed to be an effect size of mean differences between groups (e.g., Cohen’s $d$) or statistics to calculate such effect sizes (e.g., sample sizes, means, standard deviations, correlations, $F$-ratios, $t$-values). Furthermore, meta-analysis calculations require a study’s sample size to calculate sampling variance and provide weights to each effect size, therefore for a study to be included it needed to report the relevant sample sizes.

Experimental studies offer a unique opportunity to better understand intelligence scores by manipulating features of the situation to potentially influence test taker scores. While this data offers interesting insight into information processing measures, manipulations can strengthen or weaken observed relationships through the overt strength of situational manipulations or by creating dichotomies that do not exist naturally (Mesmer-Magnus & DeChurch, 2009). For these reasons, scores from experimental studies were only included from conditions where no manipulation existed (i.e., control conditions) to avoid contamination from manipulated factors.

Several key criteria about the sample for a study were also used to determine if the study was included in the meta-analysis. To ensure that the sample reflects the general working population, only studies that included an employee sample or a sample of college students were included. Studies that solely recruited participants with psychological or cognitive disorders were not included. Since examinations of group score differences were limited to the United States, only samples reporting group score differences between African Americans and Caucasians in the United States were retained. For samples reporting on performance outcomes, the data was retained even when the data was collected outside of the United States. However, only articles written in English were included.
Coding

For each study that met the inclusion/exclusion criteria, it was coded on a set of criteria related to basic features of the study (see Appendix A for full coding sheet). Each study was coded by two independent coders. I was a coder on all studies and a research assistant acted as the second coder. All research assistants underwent a two-part training process in which they were first introduced to the theory underlying the project and then trained on best practices for coding as well as the specific coding categories for this project. Research assistants then completed a practice coding, after which they received feedback. Once coding was complete, any disagreements were resolved through consensus between coders. Since there were multiple research assistants, the configuration of coders was not the same across all effect sizes, which influenced how interrater reliability was calculated for each variable type. For categorical variables, Fleiss’ kappa was calculated and for continuous variables, intraclass coefficient was calculated. These statistics are reported below for all variables included in analyses.

Sample characteristics. Every sample was coded for key features of the study, such as author(s), publication year, publication status (e.g., journal paper, presentation, dissertation, technical manual), sample type (e.g., employees, students), demographics of the sample (e.g., sex, race, age), and country of data collection. Furthermore, features of the information processing measures at the sample level were coded, such as the name of the measure, the version of the measure, number of items in the measure, whether the measure was adapted, the administration method (i.e., computer-based or paper-and-pencil), whether the test was timed or not, and the reliability of the measure in that sample.

Effect sizes and sample sizes. Next, all relevant statistics regarding the target variables (i.e., information processing, performance, group score differences) were collected. This
included means, standard deviations, sample sizes, and effect size(s) or statistics to calculate an effect size. For the validity effect sizes (i.e., correlation coefficients), group score difference effect sizes (i.e., $d$ scores), and sample sizes, there was high agreement between coders ($ICC = 1.00$).

**Moderators.** In addition to general sample characteristics, each effect size was coded based on the three variables hypothesized to be moderators. These variables include whether the measure was from the psychometric or cognitive science approach, the number of item types in the measure, and the degree to which language knowledge is required for the measure. When possible, these variables were coded based on information provided by the study that the effect size was coded from. However, in some cases the information was not directly available in the primary source materials. In those cases, the relevant information was coded from a supplementary source and imported into the dataset at a later time. In some cases, the measures were not publicly available, making it difficult to determine the specific features of the measure. In these cases, I reached out to the author or test distributor of the measure to gather additional information (e.g., request the specific information needed or a test manual for the measure).

**Theoretical approach.** Past research (e.g., Ackerman et al., 2005; McGrew, 1997) has provided comprehensive breakdowns of information processing measures into different categories. These categories include fluid reasoning, executive attention, and working memory, as just some examples, which correspond directly to the current conceptualizations of the psychometric and cognitive science approaches. Furthermore, many test designers explicitly link their measures to specific theoretical paradigms, such as the CAS to the PASS theory, which is directly derived from the cognitive science approach (Naglieri & Das, 1997). Tables 1 and 2 map the categories of the most well-known information processing measures and their associated
theoretical approach based on past literature. Coders were given these tables to use when coding a measure as rooted in the psychometric or cognitive science approach. As new measures were encountered, they were added to this master list. The Fleiss’ kappa for theoretical approach was high at .98.

**Diversity of task types.** The diversity of task types has been recognized as an important feature of intelligence measures (e.g., Jensen & Weng, 1994), though to my knowledge no prior study has operationalized this feature yet. For the purposes of this meta-analysis, item type was operationalized as discrete strategies for assessing information processing. This means there must be a fundamental difference in the construct being assessed or the method for assessing the same construct to be considered a different task. As an example, complex span measures involve the same strategy for assessing working memory, even though each span task differs in its focal action (e.g., an operation span task requires simple math skills whereas a reading span task requires simple reading skills). In this case, the tasks are the same in that they require the test taker to maintain a certain piece of information in memory. As such, these would not be considered different task types. Alternatively, the self-ordered pointing task and the recency discrimination task both assess working memory but do so in fundamentally different ways. Therefore, they would be considered different task types. The number of task types was coded as a count of these discrete strategies for assessing information processing. The ICC for the number of task types on a test was .99.

**Language knowledge requirements.** While all measures require some degree of language knowledge to understand the measure’s instructions, measures differ in the degree that language comprehension is needed to respond to each item. Language knowledge requirements can be considered from many different perspectives, such as the syntactic complexity or readability of a
sentence. However, because most information processing measures purposefully avoid the use of language to complete items, there is a low threshold for what is considered a language requirement for information processing measures. For the purposes of this meta-analysis, there were two coding variables related to language knowledge requirements. First, all effect sizes were coded as either requiring language knowledge or not requiring language knowledge. If a measure did require language knowledge, then the percentage of items on the measure that involved language was also coded. Since not all measures report the exact number of items that contain language and/or the total number of items in a measure, each coder reported an estimated percentage. For the categorical coding of whether a measure required language knowledge or not, the Fleiss’ kappa was high at .93. The ICC for estimated percentage of language on a test was .99.

Analyses

**Correction procedures.** For the current meta-analysis, Hunter and Schmidt’s (2004) psychometric procedures for correcting effect sizes were employed. This approach offers a wide range of strategies for correcting statistical artifacts that result from the various forms of error accompanying any given effect size. There are several potential statistical artifacts that can be corrected for in a meta-analysis, most often including sampling variance and error of measurement. Sampling variance reflects how an observed correlation will randomly vary from the true value of the population due to sampling error (Hunter & Schmidt, 2004). In order to correct for this artifact, sampling variance is estimated for each effect size. The inverse of the sampling variance is then used as a weight for each effect size to correct for this artifact.

Error of measurement corresponds to the random error associated with a particular measure when its utilized. Since most measures contain some level of random error, it is
considered important to reduce this artifact to systematically improve the validity estimates included in the meta-analysis. Error can exist in both the criterion and predictor variables, though correcting for them has different implications for the resulting estimates and these implications differ based on the type of effect size being examined. When examining validity coefficients, corrections are typically only done for the measurement error in the criterion. This is often considered the most accurate estimate of the “operational validity” of the predictor variable. That is, removing the random error in the criterion is believed to provide the most accurate assessment of how well the predictor correlates with an estimate of “true” performance. While measurement error in the predictor can also be corrected for, it is often seen as an inaccurate estimate of how the predictor measure operates in actual decision-making contexts (Kuncel et al., 2004). For example, when making a hiring decision, decision makers rely on the observed score from a selection tool instead of an estimate of the “true” score for that measure. Since the main interest of the current study is to estimate how information processing measures predict performance in real world situations and is not aimed at the theoretical relationship between these constructs, estimates correcting only for measurement error in the criterion are most appropriate, which is in line with prior intelligence meta-analyses (e.g., Hülsheger et al., 2007; Salgado et al., 2003). Therefore, when estimating validity coefficients, I present correlation coefficients corrected for measurement error in the criterion.

In contrast to validity coefficients where the predictor is used to make decisions, in group difference analyses, it is often the criterion that is the focal variable for decision making. As such, correcting for measurement error in the criterion leads to estimates that do not reflect the operational nature of these measures. To understand how score differences lead to differential outcomes for groups in real world situations, it is best to evaluate estimates without correcting
for the criterion. Since the predictor in group difference analyses is group membership (e.g., a person’s race), there is often no need to correct for measurement error in the predictor because these measures contain little error. For the current study, when analyzing group differences, I only report the uncorrected group differences.

Selecting the reliability coefficients to correct for measurement error is an important step since the use of inaccurate reliability estimates can lead to imprecise estimates and improper conclusions. However, there is no consensus in the field about what is the most appropriate estimation of reliability when correcting for measurement error (e.g., Murphy & DeShon, 2000; Richardson & Norgate, 2015; Schmidt, Viswesvaran, & Ones, 2000). In many ways, most reliability coefficients have some flaw and will lead to under- or overestimations of the validity. To best fit the methods of past meta-analyses and allow my findings to be comparable to the results from other meta-analyses, I corrected for measurement error in the criterion by extracting reliability estimates from prior meta-analyses. To most accurately reflect the criterion variables in my meta-analysis, I identified the most specific criterion types in prior meta-analyses to match with my dataset. In total, I identified 11 reliability coefficients from prior meta-analyses that fit with the performance outcomes in my study. I imputed these values for each relevant effect size. The criterion categories, associated reliability coefficients, and sources of these coefficients are presented in Table 3.

There are several other types of artifacts that can be corrected for when the relevant information is available. A common artifact corrected for in prior intelligence meta-analyses is restriction of range. Restriction of range refers to situations in which the sample being used to examine a relationship between variables is not random because the sample has been previously screened on some variable. As such, the sample is not representative of the population. While
restriction of range has been corrected for in prior research, only recently formulas have been derived that more accurately address indirect restriction of range, which is the most common restriction of range that impacts intelligence measures (Hunter, Schmidt, & Le, 2006). Indirect restriction of range refers to situations in which a person is screened on a variable that is not being examined in a correlation. For example, if job applicants were selected into a job based on an interview but a study was examining the relationship between an intelligence measure and job performance, this would constitute indirect restriction of range. Unfortunately, while studies in the social sciences have begun to report more information about sample characteristics, it is still extremely difficult to obtain all of the relevant information to correct for indirect range restriction (Aguinis et al., 2011; Aytug, Rothstein, Zhou, & Kern, 2012). Others have noted that the relevant data for range restriction for fluid reasoning measures is rarely found in primary studies (Postlethwaite, 2011). In line with this finding, the majority of samples in the current meta-analysis did not provide enough information to investigate the effect of range restriction, and therefore this correction was not applied. It should be noted that not correcting for range restriction means the findings in the current study should be seen as conservative estimates, particularly compared to other meta-analyses.

**Analytic strategy.** Many of the studies included in the meta-analysis contain multiple measures of information processing, including both psychometric and cognitive science measures. For that reason, effect sizes of these measures are non-independent. There are several strategies for handling non-independent samples, including computing average effect sizes within a sample or selecting a single effect size from a study (Bowman, 2012). These strategies reduce the number of effect sizes available for a meta-analysis and, in the case of the current study, would potentially create confounds when trying to parse out the effects of the different theoretical approaches. To avoid losing data associated with each sample, it is advisable to use
multi-level analyses (Bowman, 2012; Cheung, 2014; Hox, 2010). In this approach, variance components are modeled at three different levels: the sampling variance of the effect sizes (level 1), the variance between effect sizes within a sample (level 2), and the variance between samples (level 3) (Assink & Wibbelink, 2016). Multi-level analyses are effective at accounting for the interdependency of effect sizes within a sample, addressing the issues of non-independent effects. For the current study, I used the restricted maximum likelihood estimation method to estimate the relevant parameters. Estimates were computed for a random-effect model, which means there is no assumption that there is one underlying population effect for each study. This allows for a more accurate estimate of the confidence intervals for the target effects and is more capable than fixed effect approaches in identifying moderating effects in the data (Hunter & Schmidt, 2000).

To test the hypotheses, I computed the meta-analytic estimates relevant for each hypothesis along with a 95% confidence interval using the metafor package in R, which allows for multilevel and moderator tests (Viechtbauer, 2010). Supplemental syntax provided by Assink and Wibbelink (2016) was also employed for conducting the multilevel analyses. Confidence intervals offer an estimate of the variability in the meta-analytic estimate due to sampling errors. This provides the range of values that the true score falls within with a 5% chance of error. Therefore, when testing hypotheses regarding the main effect of information processing on performance, it should be expected that the confidence interval does not include zero and is positive. Similarly, for hypotheses suggesting that information processing measures have smaller group score differences than measures of general intelligence, it would be expected that the upper bound of the confidence interval is lower than 1.00 since the typical $d$ value of intelligence measures is 1.00.
Prior to testing the moderation effects, I examined if there was a significant amount of residual heterogeneity in the data to be explained. There are several tests of residual heterogeneity that can be assessed. For the current study, I first used the $Q$-test to assess the overall heterogeneity of the main effect models and then assessed the amount of variance associated with the sampling variance to test for explainable heterogeneity. The $Q$-test provides a general assessment of heterogeneity across all three levels of variance in the model (Cochran, 1954). When the $Q$-test is significant, it indicates that there is a significant amount of heterogeneity not explained by the model. However, it does not specify at what level of the model the heterogeneity is located. To better understand if the variance is due to the sampling variance or due to within-sample or between-sample variance, it is useful to compute the amount of variance associated with the sampling variance (Kepes, McDaniel, Brannick, & Banks, 2013). When the variance due to the sampling variance is below 75%, meta-analysis guidelines suggest it is reasonable to test for moderators (Hunter & Schmidt, 2004). Based on the outcomes of these analyses, I determined if it was appropriate to test my moderation hypotheses.

Moderation is tested by examining the amount of overlap between the confidence intervals of the various levels of the moderator. When there is no overlap, there is strong evidence of moderation. When the confidence intervals do overlap, it is important to assess the degree of overlap to better understand if there is any evidence of moderation. It should be noted that because of the nascent state of the information processing literature, particularly of the cognitive sciences approach, my moderator analyses are based on small sample sizes. Therefore, tests of moderation are conservative since small sample sizes often cause confidence intervals to be very wide due to their large standard error. Nonetheless, these analyses should provide an
overview of the trends in the data and offer insight into areas where additional research is needed to improve our current knowledge of these moderators.

**Results**

Data collection resulted in two separate datasets: one for the correlation coefficients (i.e., validity data) and one for the \(d\) values (i.e., the group differences data). While there was some overlap between the samples in the two datasets, most samples were unique. The final validity dataset consisted of 100 sources of data, including articles, technical reports, and conference presentations. These sources resulted in a total of 121 samples, which had a cumulative sample size of 20,055 and 234 effect sizes. There was a large range in the observed correlations between information processing measures and performance measures reported in the literature, spanning from \(-.31\) to \(.76\). The group differences dataset included 20 data sources and 23 samples. The total sample size was 19,190 with 32 effect sizes. The \(d\) values also had a large range from \(-.14\) to \(1.09\).

**Hypothesis Testing: Main Effects**

When testing my hypotheses concerning the validity of information processing measures, I began by examining the aggregated effect size across all performance outcomes (i.e., job performance, job-oriented training performance, and academic performance). While each outcome type provides unique information regarding performance, aggregating across them offers a more comprehensive understanding of the relationship between information processing measures and performance. Furthermore, when breaking performance outcomes down into specific types, the number of samples and effect sizes decreased dramatically, making it difficult to arrive at generalizable conclusions. Guidelines for meta-analyses recommend that analyses should not be conducted when there are less than five data points within a group since the
estimates may be unstable (e.g., Arthur, Bennett, & Huffcutt, 2001). While I present analyses for job performance, job-oriented training performance, and academic performance, some of the analyses should be interpreted cautiously as the results are based on a small number of sources, samples, and/or effect sizes.

**Hypothesis 1.** The first hypothesis proposed that informing processing measures will correlate with performance outcomes. The results for this hypothesis are located in Table 4. After weighting the mean observed correlation by the sampling variance (henceforth referred to as the mean weighted correlation), the relationship between information processing measures and performance outcomes is .19. Correcting for the measurement error in the criterion increased the correlation to .23. The confidence interval for the mean corrected coefficient excludes zero (95% CI: .18, .28), implying these findings are not simply a result of sampling error. Therefore, this positive effect can be considered generalizable.

When examining the effects for each of the performance outcomes, a similar pattern of findings was found. For job performance, the mean weighted correlation is .19 and the criterion mean corrected correlation is .25 (95% CI: .16, .33). Fitting with prior conceptualizations of job performance (e.g., Dalal, 2005; Lievens et al., 2008; Ones, Dilchert, & Viswesvaran, 2012; Robinson & Bennett, 1995; Rotundo & Sackett, 2002; Viswesvaran & Ones, 2000), this group of outcomes was comprised of a wide range of measures, including supervisor ratings, work samples, objective productivity indicators, and simulations. While these various types of measures allow for a broader assessment of the job performance domain, they also approach the criterion in fundamentally different ways. While measures such as work samples and simulations typically operationalize job performance as the final results on a task, ratings capture a range of behaviors related to job duties and requirements that are inherently more subjective in nature.
Because of this difference, I also ran separate analyses on 1) subjective measures of job performance (e.g., supervisor and peer ratings) and 2) objective measures of job performance (e.g., simulation, work sample, and objective productivity measures).

The results from splitting job performance into subjective and objective measures demonstrated a fairly large difference between the correlations. For subjective job performance ratings, the mean weighted correlation is .14 and the mean corrected correlation is .19 (95% CI: .10, .28). Alternatively, for objective measures of job performance, the mean weighted correlation is .30 and the mean corrected correlation is .33 (95% CI: .18, .49). While the confidence intervals for these correlations overlap, signifying they are not categorically different from each other, the resulting trend suggests a slightly stronger relationship between information processing measures and objective measures of job performance than subjective measures of job performance. This fits with prior meta-analytic findings suggesting that objective and subjective measures of job performance are not interchangeable and therefore may have distinct relationships with other variables (Bommer, Johnson, Rich, Podsakoff, & MacKenzie, 1995).

Subjective measures are more susceptible to contamination from rater bias and insufficient observational opportunities for the rater, potentially capping how accurate they reflect the criterion (e.g., Landy & Farr, 1980). However, objective measures can also be limited as they often represent a narrow set of behaviors and outcomes and capture an employee’s maximal performance more than their typical performance (Ones et al., 2012). As such, while the current findings show a gap between the subjective and objective job performance correlations, the correlations should be considered in conjunction to understand the overall effectiveness of information processing measures in predicting job performance as a general construct.
For job-oriented training performance, the mean weighted correlation is .22 and the correlated correlation is .25 (95% CI: .11, .38). Again, the confidence interval did not include zero, demonstrating a positive relationship between information processing and job-oriented training. However, it should be noted that for both objective measures of job performance and job-oriented training performance, the confidence interval around the correlation was fairly wide. This is a result of a large standard error for both variables, which is partially due to the smaller number of samples and effect sizes for these outcomes. While this does not invalidate the findings of the current study, it does emphasize the need to interpret these findings cautiously until a larger number of samples and effect sizes are available in the literature.

Finally, academic performance also showed a positive mean weighted correlation at .19 and a mean corrected correlation of .21 (95% CI: .17, .24). Taken together, these findings demonstrate consistent evidence that information processing measures correlate with performance outcomes, though there may be some variation across the outcome types. Thus, there was support for Hypothesis 1.

**Hypothesis 2.** The second hypothesis proposed that information processing measures, as a whole, would have smaller group differences than has been traditionally found on intelligence measures, which has been estimated at a $d$ of 1.00 (e.g., Hough et al., 2001; Roth et al., 2003). Table 5 provides the results for this hypothesis. Overall, the mean weighted $d$ value is .41, less than half the size of the value typically associated with general intelligence measures. The confidence interval for this value ranged from .17 to .65, showing a wide range but not including 1.00. Therefore, Hypothesis 2 was supported.

**Hypothesis Testing: Moderators**
The subsequent hypotheses all offered potential moderators of the relationship between information processing measures and performance outcomes or of the group score differences on information processing measures. Prior to testing these moderation effects, I first tested the heterogeneity of the effect sizes in each dataset to ensure moderation tests were appropriate.

Across all performance outcomes in the validity dataset, results from the $Q$-test showed that there is significant unexplained heterogeneity in the model ($Q (233) = 866.63, p < .001$). Furthermore, breaking the variance down into its components showed that only 22.35% of the variance is attributed to the sampling variance while the remaining 77.65% of variance was explained by the within-sample and between-sample levels of the model. Based on Hunter and Schmidt’s (2004) guidelines, it was determined that there is substantial variance that could still be explained by examining potential moderators.

Next, I tested the group score differences dataset to see if there was sufficient heterogeneity to be explained by potential moderators. Results from the $Q$-test showed that there was significant unexplained heterogeneity in the model ($Q (35) = 333.16, p < .001$). Examining the variance by the different levels of the multi-level model, it was possible to see that only 4.76% of the variance was attributed to the sampling variance, meaning 95.24% of the variance was explainable by the within-sample and between-sample levels of the model. Based on this classification of the variance, it was appropriate to test moderators of the group differences.

**Hypothesis 3.** In the third hypothesis, I proposed that the theoretical approach used to develop an information processing measure would moderate the relationship between the measure and performance outcomes. Specifically, I expected to see stronger correlations when examining measures from the cognitive science approach than the psychometric approach. Findings for this hypothesis are presented in Table 6. It is important to first note that overall the
cognitive science measures had fewer samples and effect sizes than the psychometric approach. This is not surprising given the long history of research associated with the psychometric approach, but has implications for the standard errors and, consequently, the confidence intervals associated with the cognitive science approach. The confidence intervals for the cognitive science approach are generally wider than the psychometric approach, making it more difficult to find instances in which the confidence intervals do not overlap. While it is important to still interpret the confidence intervals as presented, it should be noted that the tests of this and subsequent hypotheses should be considered a conservative approach due to the smaller sample sizes. In accordance with this approach, I also offer some insights based on the general trends of the differences between the correlations but emphasize that more research is needed prior to making any conclusions regarding these trends.

The mean weighted correlation between the psychometric approach and all performance outcomes is .18 and the mean corrected correlation is .22 (95% CI: .16, .27). For the cognitive science approach, the mean weighted correlation is .23 and the mean corrected correlation is .28 (95% CI: .19, .38). While the confidence intervals of the two theoretical approaches overlap, it should be noted that the upper limit of the psychometric approach’s confidence interval does not contain the mean corrected correlation of the cognitive science approach. Thus, while there are no statistically significant differences between the correlations, there is some support that the cognitive science approach is trending towards a stronger relationship with performance outcomes than the psychometric approach.

Splitting the performance outcomes into their distinct types shows that the cognitive science approach generally still has a stronger correlation than the psychometric approach, though the confidence intervals consistently overlap. For job performance, the psychometric
approach has a mean weighted correlation of .16 and a mean corrected correlation of .21 (95% CI: .12, .30), whereas the cognitive science approach has a mean weighted correlation of .22 and a mean corrected correlation of .29 (95% CI: .15, .43). The correlations between the information processing measures with objective job performance outcomes parallel the overall job performance finding. Specifically, the psychometric approach has a mean weighted correlation of .23 and a mean corrected correlation of .25 (95% CI: .09, .41) and the cognitive science approach has a mean weighted correlation of .34 and a mean corrected correlation of .38 (95% CI: .18, .57). While the confidence intervals for both overall job performance outcomes and objective job performance outcomes do not overlap completely, there is a large amount of commonality. Interestingly, when examining only the subjective job performance outcomes, there is a reverse in correlations such that the mean weighted correlation and mean corrected correlation for the psychometric approach are .15 and .20 (95% CI: .10, .29), respectively, and .12 and .17 (95% CI: -.03, .37) for the cognitive science approach. While the confidence interval of the psychometric approach fits entirely within the confidence interval for the cognitive science approach, it is important to note the confidence interval for the cognitive science approach includes zero, meaning the positive mean correlation is possibly due to sampling variance and may not be representative of the population effect.

There was only one study in the current dataset that examined job-oriented training performance from a cognitive science approach. While there were two effect sizes in this study, the sample size is so low that it makes it difficult to make any generalizable conclusions at this time. The findings do show a similar trend to the findings for overall performance. Specifically, while the psychometric approach shows a large mean weighted correlation of .22 and mean corrected correlation of .24 (95% CI: .10, .38), the cognitive science approach has a higher mean
weighted correlation of .32 and mean corrected correlation of .39 (95% CI: .01, .77). Due to the extremely large standard error for the cognitive approach mean correlation, the confidence interval for the psychometric approach is enclosed within the confidence interval for the cognitive sciences approach.

Academic performance has a mean weighted correlation of .18 and mean corrected correlation of .20 (95% CI: .16, .24) for the psychometric approach and a mean weighted correlation of .23 and mean corrected correlation of .25 (95% CI: .15, .35) for the cognitive science approach. While the lower bounds of the confidence intervals were the same, the confidence interval for the psychometric approach did not include the mean corrected correlation of the cognitive science approach. This suggests that while there is a large amount of overlap in the confidence intervals, the cognitive science approach is trending as a stronger effect than the psychometric approach.

The findings across all performance outcomes do not fully support Hypothesis 3, although they do demonstrate a general trend that the cognitive science approach’s effect size is slightly stronger than the effect size for the psychometric approach. These results should be interpreted carefully as the lack of support for the hypothesis partially stems from the large standard errors for the mean correlations. As more data is collected in this area, these findings should be reevaluated.

**Hypothesis 4.** Next, I tested whether measures from the cognitive science approach have smaller group differences than measures from the psychometric approach (see Table 7 for results). The cognitive science approach did have a smaller $d$ value (.40) than the psychometric approach (.46), though the confidence interval for the psychometric approach (95% CI: .27, .64)
fit entirely within the confidence interval for the cognitive science approach (95% CI: .11, .69). Overall, Hypothesis 4 was not supported but does have some promising early evidence.

**Hypothesis 5.** The next two moderator hypotheses examined whether an increase in the diversity of tasks within a measure strengthened the relationship between information processing measures and performance outcomes. Separate hypotheses were proposed for the psychometric approach (Hypothesis 5) and the cognitive science approach (Hypothesis 6) since diversity of tasks has different implications for each theoretical approach. Both theoretical approaches displayed a similar tendency for measures to mostly use only one type of task. For the psychometric approach, 116 (71.2%) effect sizes were based on a measure with only one task type. Of the other effect sizes, 16 had two task types (9.8%), nine had three task types (5.5%), and 22 had four task types (13.5%). A similar pattern surfaced for the cognitive science approach as 44 effect sizes had only one task type (44%), 10 with two task types (14.1%), 10 with three task types (14.1%), five with six task types (7%), and two with eight task types (2.8%). Since nearly half of the effect sizes only contained one task type, this created a highly skewed distribution for both theoretical approaches and limited the number of effect sizes representing measures with more than one task type. Because of these uneven distributions of tasks, I collapsed effect sizes into two categories: 1) measures with no diversity in task types and 2) measures with diversity in task types. Therefore, my analyses for both Hypothesis 5 and 6 are conducted with diversity of task types as a dichotomous moderator.

In testing Hypothesis 5, I examined the effect of task diversity on performance outcomes for the psychometric approach only (see Table 8 for results). Across all performance outcomes, I found that measures with no diversity had a mean weighted correlation of .21 and a mean corrected correlation of .24 (95% CI: .17, .31). Counter to my hypothesis, these correlations were
higher than the mean weighted correlation of .13 and mean corrected correlation of .16 (95% CI: .08, .24) for measures with diverse task types. While the confidence intervals overlap for both groups, only half of the confidence interval for measures with no task diversity corresponds with the confidence interval for measures with task diversity. Therefore, there is some evidence that measures with no diverse task types have a slightly higher correlation than measures with diverse task types.

A different pattern of findings appears when examining the various job performance outcomes. For all job performance outcomes, there is little difference between measures with no task diversity and those with task diversity. Specifically, for measures with no task type diversity, the mean weighted correlation is .16 and the mean corrected correlation is .21 (95% CI: .11, .32), which is identical to the mean weighted correlation of .16 and mean corrected correlation of .21 (95% CI: .05, .36) for measures with diverse task types. The findings for only the subjective job performance measures shows a different pattern in which the mean weighted correlation of .14 and mean corrected correlation of .17 (95% CI: .05, .29) for measures with no task diversity are slightly lower than the mean weighted correlation of .16 and the mean corrected correlation of .22 (95% CI: .09, .35) for measures with task diversity. It should be noted that the confidence interval for measures with no task diversity is contained entirely within the confidence interval for measures with task diversity. For objective measures of job performance, there was only one sample and effect size for measures with diverse task types. As such, the findings presented here should not be taken to reflect the true population effect until more data is collected. The weighted and mean corrected correlations for measures with no task diversity are .23 and .25 (95% CI: .03, .48), respectively, which are larger than the weighted and mean corrected correlations for the single effect size of measures with diverse task types, at .20
and .22 (95% CI: -.43, .88). Taken together, there seems to be little difference in the job performance validity for measures with and without diverse task types, though there may be a slight benefit to measures with diverse task types when examining subjective job performance outcomes.

For job-oriented training performance, measures with no task diversity have a mean weighted correlation of .25 and mean corrected correlation of .27 (95% CI: .13, .42). This is much higher than the mean weighted correlation of .06 and mean corrected correlation of .08 (95% CI: -.12, .27) for measures with diverse task types. Despite the .19 difference between these mean correlations, the confidence intervals still overlap, though the confidence interval for measures without diverse task types does not include the mean correlation for measures with diverse task types. However, it should be noted that 12 of the 13 effect sizes for measures with diverse task types came from only two samples, both of which were from the same paper. Additional research should further investigate these types of measures prior to making any conclusions regarding their validity.

Academic performance also showed a relatively large difference between the mean correlations. For measures with no diversity in tasks, the mean weighted correlation is .20 and the mean corrected correlation is .22 (95% CI: .17, .27). This is compared to the lower mean weighted correlation of .12 and mean corrected correlation of .13 (95% CI: .04, .21) for measures with task diversity. Again, while the confidence intervals for both groups overlap, there is a clear trend suggesting the measures with no task diversity have a stronger correlation with academic performance.

Overall, the evidence does not support Hypothesis 5. First, the confidence intervals for measures with and without task diversity overlapped across outcome types, though to varying
degrees. However, even when there was some evidence of a general trend, it was measures with no task diversity that had a stronger relationship with performance outcomes, counter to my original hypothesis.

**Hypothesis 6.** The sixth hypothesis proposed that measures with task diversity would lead to a stronger relationship between cognitive science measures of information processing than measures without task diversity. As with Hypothesis 5, I tested this hypothesis by splitting measures into those with and without task diversity for the moderator. The results from these analyses are presented in Table 9. Examining the effect across all performance outcomes, measures with no task diversity have a mean weighted correlation of .20 and mean corrected correlation of .26 (95% CI: .13, .40), whereas measures with task diversity have a mean weighted correlation of .28 and a mean corrected correlation of .32 (95% CI: .18, .46). These confidence intervals overlap considerably as the confidence interval for measures with no task diversity is subsumed within the confidence interval for measures with task diversity. Therefore, while measures with diverse task types have a slighter higher mean correlation with performance outcomes than measures without diverse task types, this difference should not be generalized to the population effect size.

Across all job performance measures, measures with no task diversity have a mean weighted correlation of .20 and a mean corrected correlation of .27 (95% CI: .08, .45). In contrast, measures with task diversity have a mean weighted correlation of .28 and a mean corrected correlation of .33 (95% CI: .12, .53). Similarly, for subjective measures of job performance, measures with no task diversity have a mean weighted correlation of .11 and a mean corrected correlation of .15 (95% CI: -.12, .42), whereas measures with diverse task types have a mean weighted correlation of .23 and a mean corrected correlation of .32 (95% CI: .07,
In both cases, there is a significant amount of overlap between the confidence intervals, though some evidence pointing to measures with diverse task types having slight stronger relationships with job performance outcomes. However, the relationship between task diversity and objective job performance outcomes appears to trend in the opposite direction. For measures with no task diversity, the mean weighted correlation is .38 and the mean corrected correlation is .43 (95% CI: .29, .57). Alternatively, measures with task diversity have a mean weighted correlation of .29 and a mean corrected correlation .33 (95% CI: .18, .47). The confidence interval of measures with task diversity is entirely within the confidence interval of measures without task diversity, showing little evidence that there is a difference between the groups despite the .10 difference in mean correlations. The sample size for all job performance outcomes was low across types, therefore it is important to interpret these findings cautiously.

Unfortunately, there were no samples including a measure with no task diversity for job-oriented training performance, making it impossible to compare across groups. For academic performance, there was a small sample size, with only four samples at each level of the moderator. For measures with no task diversity, the mean weighted correlation is .22 and the mean corrected correlation is .24 (95% CI: .13, .35). For measures with task diversity, the mean weighted correlation is .24 and mean corrected correlation is .27 (95% CI: .15, .38). Again, while there was a slight trend favoring measures with task diversity, the confidence intervals overlapped extensively. As a whole, the small sample sizes for the cognitive science measures makes it difficult to arrive at any concrete conclusion regarding the effect of task diversity on the relationship with performance outcomes.

**Hypothesis 7.** The final validity hypothesis examined if information processing measures that require no language knowledge will have a stronger relationship with performance outcomes
than measures that require language knowledge. The majority of effect sizes (78.2%; $k = 183$) included an information processing measure that did not require any language knowledge. This fits with the basic nature of information processing measures as they deemphasize the need for any prior knowledge. For those measures that did require some level of language knowledge, it was often difficult to identify how much of the test required language knowledge. In several cases, it was possible to identify the number of subtests that required language knowledge, but this did not provide a specific percentage of the test that required language knowledge.

Therefore, only general estimates could be provided. Due to the low number of scales that required language knowledge and the difficulty in establishing a clear percentage of the test that required language knowledge, I decided to collapse the moderator into two groups, tests that required language and tests that did not require language. Table 10 presents the results for Hypothesis 7.

When all performance outcomes are considered, measures with no language knowledge requirement have a mean weighted correlation of .19 and a mean corrected correlation of .22 (95% CI: .17, .28). The correlations for measures with language knowledge requirements are slightly higher at .23 for the mean weighted correlation and .28 (95% CI: .18, .38) for the mean corrected correlation. The confidence intervals overlap considerably, with the confidence interval for each level of the moderator enclosing the mean correlation of the other level, suggesting little evidence of moderation. However, there is a slight trend favoring measures with a language knowledge requirement, contradicting my hypothesis.

Across job performance outcomes, measures with a language knowledge requirement again show a slight advantage over measures with no language knowledge requirement, though the confidence intervals still overlap. For measures with no language knowledge requirement, the
mean weighted correlation is .17 and the mean corrected correlation is .22 (95% CI: .14, .30), whereas for measures with a language knowledge requirement the mean weighted correlation is .24 and the mean corrected correlation is .31 (95% CI: .13, .48). Interestingly, when examining only subjective job performance outcomes, this trend reverses such that the mean weighted correlation (.14) and mean corrected correlation (.19; 95% CI: .11, .28) for measures with no language knowledge requirement are slightly stronger than the mean weighted correlation (.12) and mean corrected correlation (.17; 95% CI: -.14, .47) for measures with a language knowledge requirement. The sample size for measures with a language knowledge requirement is noticeably small ($k_{samples} = 6, k_{effects} = 11$), causing a large confidence interval that includes zero. As such, this finding should not be generalized without additional research. For objective job performance outcomes, the pattern of results showed measures requiring some language knowledge had a slightly stronger correlation than measures with no language knowledge requirement. Specifically, measures with no language knowledge requirement have a mean weighted correlation of .25 and a mean corrected correlation of .28 (95% CI: .13, .43), whereas measures with a language knowledge requirement have a mean weighted correlation of .36 and a mean corrected correlation of .40 (95% CI: .18, .62). The confidence intervals do overlap, suggesting the difference in means is not generalizable.

The sample size for job-oriented training was extremely small for measures with a language knowledge requirement, with only one sample and two effect sizes. The trend does support what was found with objective job performance outcomes, that measures with a language knowledge requirement have a stronger correlation (weighted = .32; corrected = .39; 95% CI: .01, .77) than measures with no language knowledge requirement (weighted = .22; corrected = .24; 95% CI: .10, .38), but the confidence intervals for both groups are very wide and overlap.
significantly. For academic performance, there was a slight advantage to measures with no language knowledge requirement, but again the confidence intervals overlapped. For measures with no language knowledge requirements, the mean weighted correlation is .18 and the mean corrected correlation is .20 (95% CI: .16, .25), whereas the mean weighted correlation is .20 and the mean corrected correlation is .22 (95% CI: .14, .30) for measures with a language knowledge requirement.

Hypothesis 7 was overall not supported. Across all performance outcomes, there was no definitive evidence to support differences between measures that require language knowledge and those that do not. A general trend appeared suggesting that measures with language knowledge requirements might actually be marginally better at predicting outcomes, except for subjective job performance, opposing the original hypothesis. However, the confidence intervals between the two groups consistently overlapped to a high degree, suggesting there is no generalizable effect based on the current data.

**Discussion**

Information processing measures have been in use for over a century, though there is little comprehensive knowledge about their overall utility. While the intelligence literature has mostly focused on other types of measures, such as those assessing crystallized intelligence, recent research has begun to highlight the role of information processing measures. Updated models of intelligence promoting fluid reasoning and cognitive science researchers advocating for novel methodologies for measurement have contributed to this development, though I-O psychology has only engaged in a limited amount of relevant research examining the value of these measures to organizations. The objective of the current paper was to illuminate this burgeoning area of research through a meta-analytic examination of information processing
measures, specifically focusing on their validity evidence and how African American and Caucasians perform on them.

Overview of Findings

Taken as a whole, the current findings offer a promising initial assessment of information processing measures and their functionality. In the following sections, I provide a summary of the findings for each hypothesis as well as an interpretation of them compared to prior findings in the literature. Table 11 provides an overview of the theoretical rationale, summary of findings, and key takeaways for each hypothesis.

Hypothesis 1. The first hypothesis predicted that information processing measures are related to a variety of performance outcomes, including job performance, job-oriented training performance, and academic performance. There was overall support for this hypothesis across all performance types. Measures of information processing had a consistent and positive relationship with performance outcomes. Furthermore, the evidence supports that these results were not simply due to sampling variance. These findings fit with prior theory suggesting that information processing is predictive of performance on complex tasks requiring substantial attentional resources, such as adapting to changing situations, thinking analytically, or multi-tasking (e.g., Aydin et al., 2005; Fagan & Ployhart, 2015; Kanfer & Ackerman, 2005). The fact that these measures relate to performance is an important finding but does not address their relative utility compared to other measures. While it is important for these measures to be predictive, it is also valuable to know how well they function compared to other predictors, particularly other intelligence measures.

Prior review articles have shown that the validities for intelligence measures vary widely, depending on context, occupations, and other factors (e.g., Ones et al., 2012; Schmitt & Fandre,
2008). Schmitt (2014) noted that validities between intelligence measures and job performance are typically in the .20s when uncorrected.¹ This generally fits with other reviews of meta-analytic findings that suggest uncorrected validities range from .22 \((k = 140)\) to .43 \((k = 24)\), but generally average around .30 (Bobko et al., 1999). Specific meta-analyses have found an uncorrected correlation between crystallized intelligence measures and job performance to be .23 \((k = 199)\) (Postlethwaite, 2011) and between general cognitive ability and job performance to be .22 \((k = 12)\) (Bertua et al., 2005), .23 \((k = 86)\) (Postlethwaite, 2011), and .26 \((k = 7)\) (Kuncel et al., 2004). The uncorrected mean correlation in the current study is lower than these estimates at .19, but the confidence interval extends up to .26, making it reasonably in range of many of these estimates.

It is important to note that many of the prior meta-analyses based their estimates on subjective ratings of job performance rather than objective measures. When comparing across uncorrected correlations for subjective job performance outcomes, it is clear that the estimate of .14 in the current study falls below these prior estimates. The confidence interval for the uncorrected correlation goes up to .20, bringing it well within range of these prior estimates, but still relatively weaker than expected. Subjective job performance ratings are frequently recognized as imperfect measures of performance as they can be influenced by systematic biases (Richardson & Norgate, 2015) and do not correlate well with objective measures of performance (Roth et al., 2005). In addition, there is often little agreement about what constitutes good performance for a particular job, making it difficult to accurately and fairly measure it

¹ For the purposes of comparing estimates across meta-analyses, I refer to uncorrected correlation estimates and their corresponding confidence intervals. This is to ensure that all estimates are approximately equivalent since most meta-analyses use slight variations in their correction methods and/or estimates. Prior review articles have similarly made comparisons across uncorrected correlations when evaluating meta-analytic estimates (e.g., Bobko et al., 1999; Schmitt et al., 1997; Schmitt & Fandre, 2008).
(Gottfredson, 1991). Nonetheless, subjective job performance ratings are some of the most commonly used measures, making this low correlation notable. On the other hand, the relationship between information processing measures and objective job performance outcomes was found to be strong at .30. This aligns with Roth and colleagues’ (Roth, Bobko, & McFarland, 2005) meta-analysis that found general cognitive ability had a .30 uncorrected correlation with work samples. It may be that information processing measures are better at predicting objective measures of job performance than subjective measures.

Estimates of job-oriented training have also varied in past meta-analyses, with one review paper reporting estimates ranging from .23 (k = 58) for verbal ability measures to .29 (k = 223) for general mental ability measures (Schmitt & Fandre, 2008) while others have reported a correlation of .38 (k = 114) for crystallized intelligence measures (Postlethwaite, 2011). Colquitt and colleagues found that measures of general cognitive ability had an uncorrected correlation of .32 with both skill acquisition (k = 17) and transfer of training (k = 3). In the current study, the uncorrected estimate for the correlation between information processing measures and job-oriented training outcomes was .22. Much like job performance outcomes, this estimate is on the lower side of prior estimates though remains in the same general range. Furthermore, the confidence interval for information processing measures has an upper limit of .34, which contains many of the prior meta-analytic estimates.

For the final performance outcome, academic performance, prior meta-analyses have mostly examined the relationship between standardized admission tests and GPA. These tests, which include the SAT, GRE, and GMAT, are often considered measures of crystallized intelligence. An overview paper reported meta-analytic estimates for such standardized tests used in graduate school admissions ranging from .27 (k = 1,231) for the GRE to .45 (k = 22) for the
Pharmacy College Admission Test (Kuncel & Hezlett, 2007). Similarly, the Miller Analogies Test, a measure of general cognitive ability relying on a person’s vocabulary knowledge, has been found to have a .27 ($k = 70$) uncorrected correlation with overall graduate school GPA (Kuncel et al., 2004). The SAT has been found to have an uncorrected correlation with cumulative undergraduate GPA at .33 ($k = 41$) (Berry & Sackett, 2009). A meta-analysis of a wider range of intelligence measures found crystallized intelligence measures, including but not limited to admissions tests, had an uncorrected correlation of .36 ($k = 139$) while measures of general cognitive ability had a correlation of .42 ($k = 78$) (Postlethwaite, 2011). In the current study, the uncorrected relationship with information processing measures was .19 with an upper limit of .22, making it the furthest from past validities out of the three performance outcomes.

By their nature, measures of academic achievement, such as GPA, reflect a person’s crystallized knowledge as these outcomes are most often assessed by tests of declarative knowledge. A student’s prior knowledge would then reasonably act as an initial baseline for future growth in knowledge. This makes the stronger relationship between crystallized intelligence measures and measures of general cognitive ability, which typically also contain knowledge components, with academic achievement reasonable. However, it does not preclude the importance of information processing measures. Across all three performance outcomes, the findings show that while information processing measures typically have lower uncorrected correlations compared to previous meta-analytic estimates for intelligence measures, the validity coefficients associated with information processing measures also demonstrate valuable validity evidence. These estimates are higher than estimates associated with other individual difference measures such as personality measures, which typically do not exceed .20 (Schmitt, 2014). Moreover, information processing is considered a broader construct than crystallized knowledge.
as it is the root of learning new skills and knowledge (Ackerman, 1996) and therefore may still provide explanatory power above and beyond commonly used intelligence measures.

**Hypothesis 2.** In conjunction with validity evidence, it is important to also assess data related to group score differences. Hypothesis 2 predicted that information processing measures would produce lower score differences between African Americans and Caucasians than the $d$ value of 1.00 typically reported for intelligence measures. The evidence showed group differences on information processing measures were well below this target with a $d$ value of .41 and an upper confidence limit of .65. While 1.00 is the most frequently reported value of group differences across review articles (e.g., Hough & Oswald, 2000; Ployhart & Holtz, 2008; Sackett & Wilk, 1994), there is variation in what has been reported. Narrative review articles have presented values that range between .80 and 1.00 across all intelligence measures (Murphy, Cronin, & Tam, 2003), .60 and 1.00 for specific abilities (Hough et al., 2001), and .72 for moderately complex jobs and .86 for low complexity jobs (Bobko & Roth, 2013). A meta-analysis found that the GATB, a measure of general intelligence, had a $d$ value of 1.14 ($k = 127$) with a lower confidence limit of 1.11 for a sample of employees (Berry, Cullen, & Meyer, 2014). The same study also found a $d$ value of .88 ($k = 7$) and lower limit of .82 for the SAT. In another meta-analysis of group score differences, a $d$ of 1.10 ($k = 105$) was found for general cognitive abilities across sample types (e.g., military, education, industrial) with a lower limit of 1.06 (Roth et al., 2001). This same article showed that industrial samples had a $d$ of .99 ($k = 34$) with a lower limit of .88 and that the Wonderlic, a common measure of crystallized intelligence, had a $d$ of 1.00 ($k = 3$) with a lower limit of .82. Researchers caution these $d$ values may be inflated when looking across all jobs and that more appropriate estimates should be calculated at the level of a job family or across jobs with similar complexity (Bobko & Roth, 2013). Nonetheless, the
current findings show that the $d$ value for information processing measures is consistently lower than estimates of $d$ for measures of intelligence across studies and review articles.

**Hypotheses 3 and 4.** The remaining hypotheses entailed moderation. In general, while the evidence showed some trends within the data, the confidence intervals for the various levels of each moderator overlapped to varying degrees. Therefore, there was often no evidence of full moderation. It should be noted that for all of these hypotheses, the sample sizes fluctuated across performance outcomes to a large degree, with some sample sizes being extremely small. In addition, the current dataset came from a wide variety of populations, industries, testing stakes, and so forth. This makes the dataset fairly heterogeneous, increasing the overall variance. As such, my discussion speaks more to the trends in the data that should be investigated by future primary research to help explain the potential effects.

The first two moderation hypotheses predicted that the theoretical approach used to develop an information processing measure would influence the validity coefficients (Hypothesis 3) and group score differences (Hypothesis 4). Regarding validity coefficients, a general trend showed that measures developed from the cognitive science approach had a slightly stronger correlation with performance outcomes than those developed from the psychometric approach. The largest difference came from measures of objective job performance, which had an .11 difference between validity coefficients favoring cognitive science measures. This aligns with the general content and methodologies of measures from the psychometric and cognitive science approaches. Information processing measures from the psychometric approach typically involve static stimuli whereas measures from the cognitive science approach involve behavioral based tasks in which the test taker dynamically interacts with the task (e.g., adapting to new rules being presented, learning conditional associations through trial and error). These types of measures are
akin to “samples” of the actual behaviors than “signs” of the behavior (Cronbach & Meehl, 1955) and may better reflect the content of an objective job performance measure. This may also explain why measures from the cognitive science approach displayed slightly lower group score differences between African Americans and Caucasians compared to the psychometric approach. The fact that these measures are mostly behavioral based suggests they may be less impacted by contamination in the measurement device. However, it should be noted that the reduction in group score differences for both theoretical approaches is an improvement over many other intelligence measures that demonstrate larger score differences. While the group score differences for information processing measures would still be considered moderate effect sizes (Cohen, 2016), it is important to recognize that there are other factors that can play into score differences beyond the content of the measure. For example, a person’s familiarity with item content (e.g., Arvey, 1972) and stereotype threat (e.g., Mayer & Hanges, 2003; McKay, Doverspike, Bowen-Hilton, & McKay, 2003) could cause differences in scores. Research has begun to examine the roots of group score differences dating back to a person’s childhood (e.g., Cottrell et al., 2015), including their exposure to the types of tasks presented on information processing measures (Greenfield, 1998). It is important to continue this line of research to evaluate the sources of group score differences in intelligence measures, including information processing measures.

**Hypotheses 5 and 6.** The next two hypotheses examined the impact of task diversity on validity coefficients for the psychometric approach and cognitive science approach, respectively. Interestingly there was a reverse in the pattern of results for the two theoretical approaches, with the psychometric approach having higher validities for measures with no task diversity than with task diversity and the cognitive science approach demonstrating the opposite effect. This might
be a result of the difference in meaning of task diversity across the two theoretical approaches. For the cognitive science approach, task diversity reflects different measurement approaches typically assessing different types of cognitive functions. As a result, diversity of tasks may imply a broader assessment of the construct and therefore improve validity. Alternatively, diversity of task type for the psychometric approach indicates differences in task content that is not inherently related to specific cognitive functions but attempts to diversify the method for assessing fluid reasoning. While this methodology fits with calls to reduce error related to a single task type, it may also introduce additional variance based on how the tasks vary in content, ultimately reducing validity coefficients rather than strengthening them. Nonetheless, past research has found that measures of fluid reasoning can benefit from diversity in tasks (e.g., Beauducel et al., 2007; Kretzschmar et al., 2016), highlighting the importance of investigating additional moderators that may explain how variation in task content benefits information processing measures from the psychometric approach in some situations but not others.

One important caveat to the findings for Hypothesis 6 is that objective job performance outcomes were predicted better by measures with no task diversity than with task diversity within the cognitive science approach. This is particularly interesting given the strong relationship between information processing measures and objective measures of job performance as a whole. While not originally hypothesized, this effect may reveal that tasks assessing a specific cognitive function that closely relate to the content of the objective performance measure may be better suited to predict performance than measures assessing a wider set of cognitive functions. This aligns with the ability-performance compatibility principle in I-O psychology (Judge & Kammeyer-Mueller, 2012; Schneider & Newman, 2015) and the bandwidth-fidelity dilemma in personality research (e.g., Hogan & Roberts, 1996; Schneider,
Hough, & Dunnette, 1996), which suggests that there should be an alignment between the level of specificity for both the predictors and the criteria. For example, simulation tasks that require high levels of attention (e.g., an air traffic simulator) may be better predicted by just a measure of attention than by a test battery that assesses attention and conditional association. While the sample size in the current dataset does not lend itself to such tests, this could be a valuable avenue for future research and may speak to the importance of aligning the content of the predictor measure with the criterion of interest.

**Hypothesis 7.** The final hypothesis examined the role of language knowledge requirement in predicting job performance outcomes. Generally, it was found that measures requiring language knowledge had stronger validities than measures without language knowledge requirements, which contradicted the original hypothesis. It is important to consider that the language knowledge requirement for all these measures is relatively low compared to other predictors of performance, including most intelligence measures. Therefore, these measures do not rely heavily on one’s crystallized knowledge, even when including language. Still, the use of language in information processing measures may increase alignment between the content of the predictor and criterion. The role of language is evident in many of the performance outcomes in the current study, such as academic performance where performance is typically a reflection of scores on written tests or objective job performance where simulations and work samples often take the form of written scenarios. The use of language has been found to activate different neurological pathways compared to spatial stimuli (Langdon & Warrington, 2000), meaning that measures using language may be processed differently than measures only using visual stimuli. In line with the ability-performance compatibility principle, it may be that the correspondence between the predictor and criterion could strengthen the relationship when
the predictor contains verbal stimuli. Future research should investigate if there are differences in predictive validity for information processing measures based on the content of the measure and criterion. It may be that while the use of visual stimuli helps reduce the role of prior knowledge in information processing measures, the misalignment with the criterion in some situations limits their validity.

**Theoretical and Practical Implications**

There are two major takeaways from the current research that apply to both researchers and practitioners. First, the current study provides clear evidence that information processing measures are valid predictors of performance that simultaneously reduce group score differences compared to the typical $d$ value of 1.00 seen in prior meta-analyses. As such, I-O psychologists should continue to evaluate the application of information processing measures for the purposes of research and practice. Second, there is a plethora of research questions that still need to be investigated to better understand the role of information processing in the intelligence literature and how that impacts the construct’s operationalization. This includes evaluating how theoretical approaches impact information processing measures and how variations in the content of these measures can improve their predictive validity.

Overall, the current validity and group score differences evidence for information processing measures strongly supports that they should be considered when developing and administering an assessment or assessment battery for the purpose of predicting job or academic performance. As with all measures, information processing measures are not perfect predictors of performance, but can provide useful information when the desired criterion is related to a person’s information processing abilities. It is important to recognize that the utility of information processing measures does not imply the exclusion of other intelligence measures.
Instead, information processing measures should be considered another resource available to applied psychologists, organizations, and academic institutions for predicting outcomes. In particular, information processing measures can be combined with other intelligence measures to create comprehensive batteries. Woodcock (1990) made a similar argument when describing the need for creating a collection of measures to assess fluid reasoning and crystallized intelligence. Specifically, Woodcock suggested that researchers should map intelligence measures onto existing models of intelligence to better understand how to capture an appropriate breadth of the intelligence domain and to identify potential gaps in the measurement space.

This suggestion aligns with the more recent cross-battery approach (XBA) (Flanagan & McGrew, 1997; Flanagan, Ortiz, & Alfonso, 2007), which suggests utilizing multiple intelligence measures in a given testing scenario in order to create a theory-driven and comprehensive test battery. A key emphasis in this approach is linking the needs of the situation to the measures that are utilized, something that is often not well articulated in I-O psychology (Agnello et al., 2015; Scherbaum et al., 2015). In the case of information processing measures, researchers and practitioners should evaluate if the context calls for information processing. For example, as organizations place more value on the ability for employees to multitask and solve novel problems, information processing measures can offer a critical resource for measuring applicant’s abilities during the hiring process. Some researchers have even suggested integrating cognitive functions into job analysis models to better understand how specific features of information processing underlie certain knowledge, skills and abilities and, consequently, relate to performance (Becker et al., 2015). In doing so, it will be easier for organization to determine the most appropriate information processing measures to utilize for any given job. When used in conjunction with other measures, including personality and crystallized intelligence measures,
information processing measures can be a vital resource to assess a broader array of individual differences important for performance.

Researchers not only suggest assessing a wide range of knowledge, skills, and abilities to cover more of the relevant job domain but also to minimize dependency on a limited number of predictors that may demonstrate group score differences (e.g., Ployhart & Holtz, 2008; Sackett & Ellingson, 1997; Sackett & Roth, 1996). Balancing the predictive validity of measures while simultaneously maintaining diversity in hiring and promotion contexts is a pressing matter facing many organizations (Pyburn et al., 2008). Information processing measures can offer a valuable opportunity to introduce valid measures of intelligence that limit group score differences. Past research has shown that when measures of cognitive abilities with lower group score differences are included in an assessment battery, they can improve the amount of diversity during hiring without negatively impacting the quality of hires (Wee et al., 2014).

The use of information processing measures can provide an attractive opportunity for organizations, particularly as they have the potential to create fairer hiring processes. In fact, some organizations have begun to adopt neuropsychological methodologies for assessment purposes (e.g., Gee, 2017). However, despite calls for more research on intelligence theories and measures, including information processing (e.g., Scherbaum et al., 2012), there is still a general lack of studies examining how characteristics of information processing measures influence outcomes. It is critical for I-O psychologists to remain on the frontier of emerging issues like these to ensure they can leverage their expertise and empirical rigor to guide organizations in making the best business decisions (Rotolo et al., 2018). From the current study, there are several discernable research areas that should be investigated further.
Foremost, a deeper understanding of the differences between the psychometric and cognitive science approaches and their measures is needed. Although there are distinct differences in the theoretical approaches, there is still ambiguity regarding how much they overlap (Nisbitt et al., 2012). Some researchers have argued that fluid reasoning and working memory are closely related due to a shared underlying process in the central executive function (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle et al., 1999; Kyllonen, 1996; Kyllonen & Christal, 1990), while others suggest that the constructs are correlated but not isomorphic (Ackerman et al., 2005). Continued research on the differences and similarities in how the two theoretical approaches conceptualize information processing can offer new insights into the distinct operationalizations of their respective measures. Through refining and developing the theoretical side of information processing, it will be possible to better evaluate the construct validity of these measures and understand if they are assessing distinct attributes of information processing or simply supplying unique methodologies to assess the same construct. However, there are some clear methodological features of information processing measures that may be potential sources of improvement for measures across both theoretical approaches. For example, the cognitive science approach has promoted several modern methodologies, particularly computer-based assessments, that could have appealing applications for assessing fluid reasoning. In particular, more dynamic stimuli could broaden the types of stimuli used in fluid reasoning measures. Similarly, most research still relies on older fluid reasoning measures, such as the Raven’s Matrices and Cattell Culture Fair Test. An evaluation of the benefits of contemporary fluid reasoning measures and methodologies from the cognitive science approach may help progress fluid reasoning measures. This type of cross-fertilization within and between theoretical approaches can evolve intelligence measures beyond just information processing.
measures and could be a unique opportunity as organizations look for more advanced methods of assessing applicants and employees.

Beyond the theoretical approach taken to develop the assessment, there are specific characteristics of information processing measures that should be evaluated to determine their impact on validity and group score differences, including, but not limited to, task diversity and language knowledge requirement. As the current findings suggest, test content can influence the validity of information processing measures and should be evaluated, particularly in relation to improving construct validity. Lievens and Sackett (2017) described several features of predictor measures that could be examined in relation to information processing measures. For example, they suggest that stimuli can be presented as textual, pictorial, auditory, dynamic audiovisual, and face-to-face interactive. This offers a large number of stimuli types that can be combined to address concerns about error associated with a single format. While currently most measures use text or figures, there are some measures in the cognitive sciences that use auditory stimuli (Konig et al., 2005). Lievens and Sackett also note that the level of contextualization can be manipulated. At the core of information processing measures is the use of novel or unfamiliar stimuli, however, recent research has examined using stimuli that all test takers are trained on prior to the test (e.g., Fagan & Holland, 2002, 2007, 2009). In this type of assessment, a person’s ability to learn novel information through a training session allows for more contextualized questions while maintaining a generally novel task. The key point is that information processing measures should continue to evolve, which requires deeper investigations of test content and item characteristics. The current meta-analysis suggests that the content of a test matters, making it an important area for deeper consideration. In particular, researchers should begin evaluating
how to better tailor information processing measures to specific jobs, tasks, and conditions in the workplace (Becker et al., 2015).

Although the current study demonstrated several positive attributes of information processing measures, it is also imperative to recognize their limitations, both practical and conceptual, in order to appropriately evaluate their utility in any given situation. Perhaps most notably is that some information processing measures, particularly from the cognitive science approach, require a significant amount of resources and time to administer and score. Many of these measures were developed for administration in lab or clinical settings, which allows them to be more dynamic but also requires computer testing that may involve specific hardware requirements. They also typically require proctoring to ensure test security and avoid errors made by the test taker. This is in contrast to many other intelligence measures that can be administered via paper and pencil in large group settings. Similarly, when developing any type of assessment that contains multiple tests, it can potentially increase the cost and time of administration (Ployhart & Holtz, 2008). Researchers and practitioners need to be cognizant of the balance between assessing a range of abilities while accounting for practical limitations. There should also be some consideration of how test takers may react to information processing measures. Past research has shown that test takers often rate measures using abstract item types as less face valid than measures of vocabulary and math (Smither, Reilly, Millsap, & Pearlman, 1993). The way that test takers perceive a measure can have important implications for test-taking motivation, job acceptance intentions, and even performance on the measure (Ryan & Ployhart, 2000). These potential limitations should not prevent researchers and practitioners from evaluating the practicality of administering information processing measures but should rather factor into their overall assessment of the measures’ utility.
Limitations

Despite best efforts, there are still limitations to the current meta-analysis that should be noted. Foremost, the sample size for the overall study was fairly low. Even when using multilevel analyses to capture multiple effect sizes within a sample, there was still a limited number of samples available for analysis. This was particularly noticeable when examining moderators, as several groups (e.g., measures from the cognitive science approach, psychometric measures with diverse tasks) were not well represented in the literature. Similarly, there was a lack of research examining the population norms for most of these measures. Even when some normative data exists, it is frequently based on populations that do not generalize to most I-O psychology research, such as children and clinical populations, making it impractical to correct for range restriction in the current study. While correcting observed correlations is a debated practice in general (e.g., Richardson & Norgate, 2015), several prior meta-analyses have reported corrected correlations for both measurement error and restriction of range. As such, direct comparisons between the estimates from the current study and past meta-analyses on the corrected values must be done carefully.

Another potential limitation is the process for identifying information processing measures and categorizing them based on the relevant moderators. First, it was necessary to decide how to operationalize information processing measures. In reality, many measures of intelligence include an information processing component, however, they are not strictly information processing measures. For the purposes of the current study, a blunt delineation was made around information processing measures. However, there may be measures that are on the edge of this distinction that could be useful for understanding the interplay between information processing and crystallized knowledge. This makes the current study a conservative assessment...
but could be reevaluated with a broader scope of information processing measures in the future. Furthermore, the classification of moderators was restricted into dichotomous variables even though continuous variables would be better suited for the study. For both diversity of tasks and language knowledge requirements, there was a restriction in variation that made it impractical to examine more nuanced analyses with the moderators.

Finally, it is important to note some caveats regarding meta-analysis findings in general. Meta-analyses are limited by the available data in primary studies. As many studies do not report information about all relevant methodological features and statistics, it can be difficult to appropriately aggregate and correct the data (Richardson & Norgate, 2015; Schmitt, Arnold, & Nieminen, 2010). For example, most studies do not report local reliabilities, meaning corrections for measurement error must be imputed based on external data, usually other meta-analyses. This can limit the appropriateness of a correction being made. Similarly, without information about a study’s sample or test administration, it can be difficult to know the degree of heterogeneity within a dataset. In addition, while attempts are made to overcome the file-drawer problem, it is impossible to know if all data has been included in a meta-analysis. This means there is still a potential bias in the dataset for published studies. These caveats do not mean meta-analyses and their findings are unusable but should be seen as a step in evaluating the current research on a topic and promoting more research in that area. It is important to follow up meta-analyses with comprehensive studies that have greater control over the factors of interest to reduce error variations across factors (e.g., Bobko & Roth, 2003; Bobko & Stone-Romero, 1998).

**Future Research**

There are many avenues for future research on information processing measures. In addition to the research areas discussed above, several key areas should be addressed to better
evaluate the utility of information processing measures. One primary line of research is examining how contextual factors influence the validity and group score differences for information processing measures. The current meta-analysis focused on measure level factors, but many past meta-analyses have examined how intelligence measures perform differently based on the content of the job (e.g., low vs. high complexity jobs) (Gonzalez-Mulé et al., 2014; Hülsheger et al., 2007; Hunter & Hunter, 1984; Salgado et al., 2003; Ziegler et al., 2011) or the stakes associated with the test (e.g., low vs. high testing stakes) (Postlethwaite, 2011). These types of moderators can elucidate the contexts in which information processing measures operate most effectively.

Another fruitful avenue of research is examining the explanatory power of information processing measures in comparison to other intelligence measures. For the current study, estimates of the validity for information processing measures were examined in isolation and then compared to existing meta-analytic estimates of other intelligence measures. To better understand the contribution of information processing measures to the validity of outcomes, they should be evaluated in studies that administer and evaluate multiple intelligence measures at once. This fits with recent intelligence theories that suggest when multiple measures of specific cognitive abilities are administered simultaneously, they can be equally or more predictive than general intelligence measures (e.g., Lang & Bliese, 2012; Lang et al., 2010; Van Der Maas et al., 2006). Through the use of relative weights analysis (Johnson & LeBreton, 2004), it is possible to assess the contribution of each intelligence measure for predicting an outcome. Such analyses can offer insight into the unique variance explained by information processing measures compared to other intelligence measures. It should be noted that even when measures do not explain much variance beyond other measures in a battery, they still may be useful as
independent measures when situational constraints limit the number of measures that can be administered.

A final area of research that will be critical going forward is examinations of mean score differences for groups besides African Americans and Caucasians. The information processing literature and I-O psychology can benefit from deeper examinations of several possible group score differences, such as those based on a person’s sex and age. Though some meta-analytic work has shown that men have a slight advantage on the Raven’s Matrices ($d = .33$) (Irving & Lynn, 2005), more research is needed to determine if this difference generalizes to other measures. In addition, age differences are perhaps the most well-known group difference related to information processing measures. Evidence suggests that as people grow older, there is a general slowing of the frontal lobe in executing cognitive functions and more difficulty retaining information in temporary storage (Bugg, Zook, DeLosh, Davalos, & Davis, 2006). This fits with early research by Horn and Cattell (1966) that showed fluid reasoning was negatively related to age. Beyond sex and age differences, more research is needed on differences between races not examined in the current study (e.g., Hispanic, Asian) as well as other demographics.

Furthermore, a principal feature of information processing measures is their reduced language requirements, making them appealing in international settings. Research is needed to better understand how these tests operate with a global population and examine potential sources of group score differences. For example, past research has found multilingual speakers may outperform monolingual speakers on measures of executive function (Bialystok, Craik, Klein, & Viswanathan, 2004). I-O psychologists should pay careful attention to these types of possible group differences as they relate directly to adverse impact and could have important implications for differential validity of information processing measures.
Conclusion

The importance of intelligence measures in I-O and educational psychology is undeniable as they are strong predictors of performance across a wide spectrum of contexts. As a result of their value to both researchers and practitioners, it is critical to continually improve upon the intelligence literature by clarifying intelligence as a construct and refining its measurement. Although information processing and its measurement have been a topic of discussion since the beginning of contemporary intelligence research, it has not been abundantly investigated to properly understand its full potential in predicting critical outcomes and reducing group score differences. The findings from the current study offer an initial evaluation of their utility, which hopefully contributes to the current revival of research on the topic. The utility of these measures in various contexts still needs to be explored in greater depth, but there is significant potential within these measures. Paired together with past research on intelligence and leveraging methodologies and findings from other disciplines, I-O and educational psychology have the opportunity to significantly advance knowledge about intelligence and how to best predict its outcomes.
## Appendix A: Coding Form

<table>
<thead>
<tr>
<th>Coder</th>
<th>Author(s)</th>
<th>Year</th>
<th>Article Title</th>
<th>Journal/Thesis/Dissertation/Conference</th>
<th>Summary of Article (1-2 sentences)</th>
<th>Does Study Include Measure of Information Processing?</th>
<th>Does Study Examine Information Processing Measure and 1) Relationship with Performance Outcomes, 2) Group Score Differences, or 3) Both?</th>
<th>Is Sample Non-Clinical and Over 18 Years Old?</th>
<th>Does Study Include Manipulations?</th>
</tr>
</thead>
</table>


<table>
<thead>
<tr>
<th>INCLUSION/EXCLUSION CRITERIA (CONTINUED)</th>
<th>STUDY DETAILS</th>
<th>SAMPLE CHARACTERISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>If Study Includes Manipulation, Does it</td>
<td>Does Study</td>
<td>Sample Type (Job Applicants; Job Incumbents; Undergraduate Students; Graduate Students)</td>
</tr>
<tr>
<td>Have a Control Group?</td>
<td>Include Correlations and/or Cohen’s d?</td>
<td></td>
</tr>
<tr>
<td>Does Study Include Correlations and/or</td>
<td>Does Study Meet Inclusion Criteria?</td>
<td></td>
</tr>
<tr>
<td>Cohen’s d?</td>
<td>Number of Relevant Studies in Paper (Use Separate Row for Each Study)</td>
<td></td>
</tr>
<tr>
<td>Does Study Meet Inclusion Criteria?</td>
<td>Is Intelligence a Focal Variable, Secondary Variable, or Control Variable?</td>
<td></td>
</tr>
<tr>
<td>If Study Does Not Include Correlation</td>
<td>Does Study Include Manipulations?</td>
<td></td>
</tr>
<tr>
<td>or Cohen’s d, Does it Include Statistics to</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calculate Correlation and/or Cohen’s d?</td>
<td>If Study Has Manipulation, What Is the Manipulation? Only Include Data for Control Conditions.</td>
<td></td>
</tr>
<tr>
<td>For Employees, Where Were They Recruited From (Participant Pool; Psychology Class)?</td>
<td>Number of Participants Listed as Another Race (List Race and Number if Reported)</td>
<td></td>
</tr>
<tr>
<td>For Students, Where Were They Recruited From? List All</td>
<td>Number of Female Participants</td>
<td></td>
</tr>
<tr>
<td>Were Participants Screened Prior to Participating? If So, What Were They Screened On?</td>
<td>Number of Male Participants</td>
<td></td>
</tr>
<tr>
<td>Number of African American/Black Participants</td>
<td>Number of Asian Participants</td>
<td></td>
</tr>
<tr>
<td>Number of Caucasian/White Participants</td>
<td>Number of Hispanic Participants</td>
<td></td>
</tr>
<tr>
<td>Number of Participants Listed as Another Race (List Race and Number if Reported)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Participants Where English is Their Primary Language</td>
<td>Number of Participants Where English is Not Their Primary Language</td>
<td>Age Range</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

**INFORMATION PROCESSING MEASURES**

| Name of Information Processing Measure (Use Separate Row for Each Measure) | Citation for Measure | Theoretical Approach for Measure (Psychometric or Cognitive Science) | Version of the Measure and/or Year it Was Published | Original or Adapted/ Modified Version of Measure | If Adapted, How? | Number of Test Items | Number of Dimensions on Test | Name of Each Dimension | Number of Task Types | List Name of All Task Types Reported | Provide Brief Description of Task Types Reported (1-2 Sentences) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|

116
<table>
<thead>
<tr>
<th>INFORMATION PROCESSING MEASURES (CONTINUED)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Do Items Include Language?</strong></td>
</tr>
<tr>
<td><strong>Dependent Variables</strong></td>
</tr>
<tr>
<td>Name or Description of Dependent Variable (Use Separate Row for Each Measure)</td>
</tr>
<tr>
<td>DEPENDENT VARIABLES (CONTINUED)</td>
</tr>
<tr>
<td>--------------------------------</td>
</tr>
<tr>
<td>Was There a Time Lag between Information Processing Measure and DV?</td>
</tr>
<tr>
<td>If Time Lag, How Long?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>If d not Reported, Provide All Relevant Statistics</td>
</tr>
</tbody>
</table>
Table 1

*Representative Sample of Psychometric Information Processing Measures*

<table>
<thead>
<tr>
<th>Psychometric Measures</th>
<th>Measure Citation</th>
<th>Sample Citations Reporting on Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability Processing of Information and Learning Battery – Concept Formation Test</td>
<td>Taylor, 1997</td>
<td>Mashu, 2015; Pretorius, 2010</td>
</tr>
<tr>
<td>Baddeley Reasoning Test</td>
<td>Baddeley, 1968</td>
<td>Chamorro-Premuzic &amp; Furnham, 2008; Furnham, 2012; Furnham et al., 2006</td>
</tr>
<tr>
<td>Berlin Test of Intelligence Structure – Fluid Subtest</td>
<td>Jäger et al., 2006</td>
<td>Kunina Wilhelm, Formazin, Jonkmann, &amp; Schroeders, 2007</td>
</tr>
<tr>
<td>Cattell Culture Fair Intelligence Test</td>
<td>Cattell, 1973</td>
<td>Belhekar, 2017; Côté &amp; Miners, 2006; Domino, 1964; Durso et al., 2016; Lowery et al., 2004; Naderi, Abdullah, Hamid, &amp; Sharir, 2009</td>
</tr>
<tr>
<td>D-48/D-70 Test</td>
<td>Black, 1961</td>
<td>Domino, 1964; Domino, 2000; McLaurin, Pendergraaff, &amp; Kennedy, 1973</td>
</tr>
<tr>
<td>Intelligenz-Struktur-Test</td>
<td>Amthauer, Brocke, Liepmann, &amp; Beauducel, 2001</td>
<td>Bühner et al., 2006; König, Bühner, &amp; Mürling, 2005; Krumm et al., 2012</td>
</tr>
<tr>
<td>Intelligence Structure Battery S2</td>
<td>Arendasy et al., 2005</td>
<td>Sommer et al., 2006</td>
</tr>
<tr>
<td>Kaufman Brief Intelligence Test</td>
<td>Kaufman &amp; Kaufman, 1997</td>
<td>Bradshaw, 2008</td>
</tr>
<tr>
<td>Primary Mental Abilities Test – Inductive Reasoning Subtest</td>
<td>Thurstone, 1938</td>
<td>Burton &amp; Dowling, 2010; Chamorro-Premuzic et al., 2009</td>
</tr>
<tr>
<td>Measure Name</td>
<td>Measure Citation</td>
<td>Sample Citations Reporting on Measure</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------</td>
<td>--------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>SRA Pictorial Reasoning Test</td>
<td>Science Research Associates, 1967</td>
<td>Fox &amp; Lefkowitz, 1974; Kirkpatrick et al., 1968; Lefkowitz, 1972</td>
</tr>
<tr>
<td>TRASI</td>
<td>Rubio &amp; Santacreu, 2003</td>
<td>Colom, Martínez-Molina, Shih, &amp; Santacreu, 2010</td>
</tr>
</tbody>
</table>
**Table 2**

**Representative Sample of Cognitive Science Information Processing Measures**

<table>
<thead>
<tr>
<th>Cognitive Science Measures</th>
<th>Measure Citation</th>
<th>Sample Citations Reporting on Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>CogScreen - AE Working Memory Subtest</td>
<td>Kay, 1995</td>
<td>Taylor, O’Hara, Mumenthaler, Rosen, &amp; Yesavage, 2005</td>
</tr>
<tr>
<td>Dorsolateral Prefrontal Cognitive Ability Battery</td>
<td>Higgins et al., 2007</td>
<td>Higgins et al., 2007; Sabet et al., 2013</td>
</tr>
<tr>
<td>Executive Attention Battery</td>
<td>Bosco et al., 2015</td>
<td>Bosco et al., 2015</td>
</tr>
<tr>
<td>Operation Span Task</td>
<td>Turner &amp; Engle, 1989</td>
<td>Bergersen &amp; Gustafsson, 2011; Durso, 2006; Hambrick et al., 2010; Periman, 2016; Rohde &amp; Thompson, 2007</td>
</tr>
<tr>
<td>Reading Span Task</td>
<td>Daneman &amp; Carpenter, 1980</td>
<td>Bergersen &amp; Gustafsson, 2011; Durso, 2016; König et al., 2005; Perlow et al., 1997</td>
</tr>
<tr>
<td>Test Battery for Attentional Performance</td>
<td>Zimmermann &amp; Fimm, 1993</td>
<td>Bühner et al., 2006; König et al., 2005</td>
</tr>
<tr>
<td>Tower of London/Hanoi</td>
<td>Shallice, 1982</td>
<td>Culbertson et al., 2013</td>
</tr>
<tr>
<td>Verbal Working Memory Task</td>
<td>Salthouse, 1992</td>
<td>Perlow et al., 1997</td>
</tr>
<tr>
<td>Wisconsin Card Sorting Test</td>
<td>Berg, 1948</td>
<td>Culbertson et al., 2013; Hidlebrand, 1996</td>
</tr>
<tr>
<td>Working Memory Test Battery</td>
<td>Oberauer, Süß, Wilhelm, &amp; Wittman, 2003</td>
<td>Bühner et al., 2006; König et al., 2005; Krumm et al., 2012</td>
</tr>
</tbody>
</table>
Table 3

*Reliability Coefficients Used for Correcting Measurement Error in the Criterion*

<table>
<thead>
<tr>
<th>Criterion Type</th>
<th>Reliability Coefficient</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Job Performance Rated by Supervisor</td>
<td>.52</td>
<td>Rothstein, 1990; Salgado et al., 2003; Salgado &amp; Tauriz, 2014; Viswesvaran, Ones, &amp; Schmidt, 1996</td>
</tr>
<tr>
<td>Job Performance Rated by Peer</td>
<td>.42</td>
<td>Viswesvaran et al., 1996</td>
</tr>
<tr>
<td>Objective Task Performance (e.g., Work Samples; Simulations)</td>
<td>.80</td>
<td>McKay &amp; McDaniel, 2006; Roth et al., 2003</td>
</tr>
<tr>
<td>Counterproductive Workplace Behaviors</td>
<td>.82</td>
<td>Gonzalez-Mulé et al., 2014</td>
</tr>
<tr>
<td>Organizational Citizenship Behaviors</td>
<td>.89</td>
<td>Gonzalez-Mulé et al., 2014</td>
</tr>
<tr>
<td>Training Course Assessment (e.g., Grade in Training Class)</td>
<td>.90</td>
<td>Brown, Le, &amp; Schmidt, 2006</td>
</tr>
<tr>
<td>Training Evaluation (e.g., Instructor Rating)</td>
<td>.56</td>
<td>Salgado et al., 2003</td>
</tr>
<tr>
<td>Transfer of Skill Test</td>
<td>.75</td>
<td>Alliger, Tannenbaum, Bennett, Traver, &amp; Shotland, 1997</td>
</tr>
<tr>
<td>Retention of Skill Test</td>
<td>.53</td>
<td>Alliger et al., 1997</td>
</tr>
<tr>
<td>GPA</td>
<td>.83</td>
<td>Kuncel et al., 2001</td>
</tr>
</tbody>
</table>
### Table 4

**Information Processing Measures and Performance Outcomes**

<table>
<thead>
<tr>
<th></th>
<th>$k_{\text{sources/samples/effects}}$</th>
<th>$n$</th>
<th>$r_w$</th>
<th>$\text{se}_r$</th>
<th>95% CI $r$</th>
<th>$r_c$</th>
<th>$\text{se}_c$</th>
<th>95% CI $r_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall performance</td>
<td>100/121/234</td>
<td>36,658</td>
<td>.19</td>
<td>.02</td>
<td>.15</td>
<td>.24</td>
<td>.23</td>
<td>.03</td>
</tr>
<tr>
<td>outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outcomes (overall)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjective job</td>
<td>31/39/75</td>
<td>13,437</td>
<td>.14</td>
<td>.03</td>
<td>.07</td>
<td>.20</td>
<td>.19</td>
<td>.05</td>
</tr>
<tr>
<td>performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Objective job</td>
<td>18/21/47</td>
<td>6,003</td>
<td>.30</td>
<td>.07</td>
<td>.16</td>
<td>.43</td>
<td>.33</td>
<td>.08</td>
</tr>
<tr>
<td>performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job-oriented training</td>
<td>16/18/48</td>
<td>8,629</td>
<td>.22</td>
<td>.06</td>
<td>.10</td>
<td>.34</td>
<td>.25</td>
<td>.07</td>
</tr>
<tr>
<td>performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic performance</td>
<td>43/49/64</td>
<td>8,589</td>
<td>.19</td>
<td>.02</td>
<td>.15</td>
<td>.22</td>
<td>.21</td>
<td>.02</td>
</tr>
<tr>
<td>outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. $k =$ number of sources of data/independent study samples/effect sizes; $n =$ cumulative total sample size for effect sizes; $r_{\text{weighted}} =$ inverse sampling variance-weighted, uncorrected mean correlation; $\text{se}_r =$ standard error of weighted, uncorrected mean correlation; $r_c =$ mean correlation corrected for measurement error in the criterion; $\text{se}_c =$ standard error of correlation corrected for measurement error in the criterion; 95% CI =$ the lower limit (LL) and upper limit (UL) of the confidence interval around the correlation corrected for measurement error in the criterion.*
Table 5

**Group Score Differences**

<table>
<thead>
<tr>
<th></th>
<th>k&lt;sub&gt;sources/samples/effects&lt;/sub&gt;</th>
<th>n</th>
<th>d&lt;sub&gt;weighted&lt;/sub&gt;</th>
<th>se&lt;sub&gt;d&lt;/sub&gt;</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LL</td>
</tr>
<tr>
<td>Overall effect</td>
<td>20/23/32</td>
<td>26,092</td>
<td>.41</td>
<td>.12</td>
<td>.17</td>
</tr>
</tbody>
</table>

*Note. k = number of sources of data/independent study samples/effect sizes; n = cumulative total sample size for effect sizes; d<sub>weighted</sub> = inverse sampling variance-weighted, uncorrected mean d value; se<sub>d</sub> = standard error of weighted, uncorrected mean d value; 95% CI = the lower limit (LL) and upper limit (UL) of the confidence interval around the weighted, uncorrected d value.*
Table 6

**Theoretical Approach and Performance Outcomes**

<table>
<thead>
<tr>
<th>Overall performance outcomes</th>
<th>ksources/samples/effects</th>
<th>n</th>
<th>rw</th>
<th>ser</th>
<th>95% Clrw</th>
<th>rc</th>
<th>sec</th>
<th>95% Clrc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychometric</td>
<td>85/99/163</td>
<td>26,856</td>
<td>.18</td>
<td>.02</td>
<td>.14</td>
<td>.23</td>
<td>.03</td>
<td>.16</td>
</tr>
<tr>
<td>Cognitive science</td>
<td>20/27/71</td>
<td>9,802</td>
<td>.23</td>
<td>.04</td>
<td>.14</td>
<td>.31</td>
<td>.28</td>
<td>.05</td>
</tr>
<tr>
<td>Job performance outcomes (overall)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychometric</td>
<td>35/41/66</td>
<td>11,044</td>
<td>.16</td>
<td>.04</td>
<td>.09</td>
<td>.23</td>
<td>.21</td>
<td>.04</td>
</tr>
<tr>
<td>Cognitive science</td>
<td>14/19/56</td>
<td>8,396</td>
<td>.22</td>
<td>.06</td>
<td>.10</td>
<td>.34</td>
<td>.29</td>
<td>.07</td>
</tr>
<tr>
<td>Subjective job performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychometric</td>
<td>27/31/51</td>
<td>8,984</td>
<td>.15</td>
<td>.03</td>
<td>.08</td>
<td>.21</td>
<td>.20</td>
<td>.05</td>
</tr>
<tr>
<td>Cognitive science</td>
<td>6/10/24</td>
<td>4,453</td>
<td>.12</td>
<td>.08</td>
<td>-.03</td>
<td>.27</td>
<td>.17</td>
<td>.10</td>
</tr>
<tr>
<td>Objective job performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychometric</td>
<td>11/12/15</td>
<td>2,060</td>
<td>.23</td>
<td>.07</td>
<td>.08</td>
<td>.37</td>
<td>.25</td>
<td>.08</td>
</tr>
<tr>
<td>Cognitive science</td>
<td>9/11/32</td>
<td>3,943</td>
<td>.34</td>
<td>.09</td>
<td>.16</td>
<td>.51</td>
<td>.38</td>
<td>.10</td>
</tr>
<tr>
<td>Job-oriented training performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychometric</td>
<td>15/17/46</td>
<td>8,445</td>
<td>.22</td>
<td>.06</td>
<td>.09</td>
<td>.34</td>
<td>.24</td>
<td>.07</td>
</tr>
<tr>
<td>Cognitive science</td>
<td>1/1/2</td>
<td>184</td>
<td>.17</td>
<td>-.02</td>
<td>.65</td>
<td>.39</td>
<td>.19</td>
<td>.01</td>
</tr>
<tr>
<td>Academic performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychometric</td>
<td>38/43/51</td>
<td>7,367</td>
<td>.18</td>
<td>.02</td>
<td>.14</td>
<td>.22</td>
<td>.20</td>
<td>.02</td>
</tr>
<tr>
<td>Cognitive science</td>
<td>6/7/13</td>
<td>1,222</td>
<td>.23</td>
<td>.05</td>
<td>.14</td>
<td>.32</td>
<td>.25</td>
<td>.05</td>
</tr>
</tbody>
</table>

Note. k = number of sources of data/independent study samples/effect sizes; n = cumulative total sample size for effect sizes; rw = weighted inverse sampling variance-weighted, uncorrected mean correlation; ser = standard error of weighted, uncorrected mean correlation; rc = mean correlation corrected for measurement error in the criterion; sec = standard error of correlation corrected for measurement error in the criterion; 95% CI = the lower limit (LL) and upper limit (UL) of the confidence interval around the correlation corrected for measurement error in the criterion.
Table 7

*Group Score Differences by Theoretical Approach*

<table>
<thead>
<tr>
<th></th>
<th>$k_{sources/samples/effect}$</th>
<th>$n$</th>
<th>$d_{weighted}$</th>
<th>$se_d$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychometric</td>
<td>14/14/15</td>
<td>3,844</td>
<td>.46</td>
<td>.09</td>
<td>.27</td>
</tr>
<tr>
<td>Cognitive science</td>
<td>7/10/17</td>
<td>22,248</td>
<td>.40</td>
<td>.14</td>
<td>.11</td>
</tr>
</tbody>
</table>

*Note. $k =$ number of sources of data/independent study samples/effect sizes; $n =$ cumulative total sample size for effect sizes; $d_{weighted} =$ inverse sampling variance-weighted, uncorrected mean $d$ value; $se_d =$ standard error of weighted, uncorrected mean $d$ value; 95% CI = the lower limit (LL) and upper limit (UL) of the confidence interval around the weighted, uncorrected $d$ value.*
Table 8

Task Diversity Within the Psychometric Approach and Performance Outcomes

<table>
<thead>
<tr>
<th></th>
<th>$k_{sources/samples/effects}$</th>
<th>n</th>
<th>$r_w$</th>
<th>se$_r$</th>
<th>95% CI</th>
<th>$r_c$</th>
<th>se$_{r_c}$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall performance outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>66/74/116</td>
<td>19,028</td>
<td>.21</td>
<td>.03</td>
<td>.15</td>
<td>.26</td>
<td>.24</td>
<td>.03</td>
</tr>
<tr>
<td>Task diversity</td>
<td>21/27/47</td>
<td>7,828</td>
<td>.13</td>
<td>.04</td>
<td>.06</td>
<td>.20</td>
<td>.16</td>
<td>.04</td>
</tr>
<tr>
<td><strong>Job performance outcomes (overall)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>27/29/46</td>
<td>6,529</td>
<td>.16</td>
<td>.04</td>
<td>.08</td>
<td>.25</td>
<td>.21</td>
<td>.05</td>
</tr>
<tr>
<td>Task diversity</td>
<td>8/12/20</td>
<td>4,515</td>
<td>.16</td>
<td>.06</td>
<td>.04</td>
<td>.28</td>
<td>.21</td>
<td>.08</td>
</tr>
<tr>
<td><strong>Subjective job performance outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>20/20/32</td>
<td>4,598</td>
<td>.14</td>
<td>.04</td>
<td>.05</td>
<td>.23</td>
<td>.19</td>
<td>.06</td>
</tr>
<tr>
<td>Task diversity</td>
<td>7/11/19</td>
<td>4,386</td>
<td>.16</td>
<td>.05</td>
<td>.06</td>
<td>.26</td>
<td>.21</td>
<td>.07</td>
</tr>
<tr>
<td><strong>Objective job performance outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>10/11/14</td>
<td>1,931</td>
<td>.23</td>
<td>.09</td>
<td>.02</td>
<td>.43</td>
<td>.25</td>
<td>.10</td>
</tr>
<tr>
<td>Task diversity</td>
<td>1/1/1</td>
<td>129</td>
<td>.20</td>
<td>.27</td>
<td>-.39</td>
<td>.79</td>
<td>.22</td>
<td>.30</td>
</tr>
<tr>
<td><strong>Job-oriented training performance outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>13/14/33</td>
<td>6,827</td>
<td>.25</td>
<td>.06</td>
<td>.13</td>
<td>.38</td>
<td>.27</td>
<td>.07</td>
</tr>
<tr>
<td>Task diversity</td>
<td>2/3/13</td>
<td>1,618</td>
<td>.06</td>
<td>.09</td>
<td>-.11</td>
<td>.24</td>
<td>.08</td>
<td>.10</td>
</tr>
<tr>
<td><strong>Academic performance outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>28/33/37</td>
<td>5,672</td>
<td>.20</td>
<td>.02</td>
<td>.16</td>
<td>.24</td>
<td>.22</td>
<td>.02</td>
</tr>
<tr>
<td>Task diversity</td>
<td>12/12/14</td>
<td>1,695</td>
<td>.12</td>
<td>.04</td>
<td>.04</td>
<td>.19</td>
<td>.13</td>
<td>.04</td>
</tr>
</tbody>
</table>

*Note. $k =$ number of sources of data/independent study samples/effect sizes; $n =$ cumulative total sample size for effect sizes; $r_{weighted} =$ inverse sampling variance-weighted, uncorrected mean correlation; $se_r =$ standard error of weighted, uncorrected mean correlation; $r_c =$ mean correlation corrected for measurement error in the criterion; $se_{r_c} =$ standard error of correlation corrected for measurement error in the criterion; 95% CI =$ the lower limit (LL) and upper limit (UL) of the confidence interval around the correlation corrected for measurement error in the criterion.*
Table 9

Task Diversity Within the Cognitive Sciences Approach and Performance Outcomes

<table>
<thead>
<tr>
<th></th>
<th>$k_{\text{sources/samples/effects}}$</th>
<th>$n$</th>
<th>$r_w$</th>
<th>$\text{se}_r$</th>
<th>95% CI $r_c$</th>
<th>$r_c$</th>
<th>$\text{se}_{r_c}$</th>
<th>95% CI $\text{se}_{r_c}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LL</td>
<td>UL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>14/15/44</td>
<td>6,471</td>
<td>.20</td>
<td>.07</td>
<td>.07</td>
<td>.33</td>
<td>.26</td>
<td>.07</td>
</tr>
<tr>
<td>Task diversity</td>
<td>8/14/27</td>
<td>3,331</td>
<td>.28</td>
<td>.06</td>
<td>.16</td>
<td>.39</td>
<td>.32</td>
<td>.07</td>
</tr>
<tr>
<td>Job performance outcomes (overall)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>10/11/35</td>
<td>5,872</td>
<td>.20</td>
<td>.09</td>
<td>.03</td>
<td>.37</td>
<td>.27</td>
<td>.09</td>
</tr>
<tr>
<td>Task diversity</td>
<td>5/9/21</td>
<td>2,524</td>
<td>.28</td>
<td>.09</td>
<td>.11</td>
<td>.45</td>
<td>.33</td>
<td>.10</td>
</tr>
<tr>
<td>Subjective job performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>4/5/19</td>
<td>33,991</td>
<td>.11</td>
<td>.09</td>
<td>-.08</td>
<td>.30</td>
<td>.15</td>
<td>.13</td>
</tr>
<tr>
<td>Task diversity</td>
<td>2/5/5</td>
<td>462</td>
<td>.23</td>
<td>.09</td>
<td>.05</td>
<td>.41</td>
<td>.32</td>
<td>.12</td>
</tr>
<tr>
<td>Objective job performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>6/6/16</td>
<td>1,881</td>
<td>.38</td>
<td>.06</td>
<td>.26</td>
<td>.51</td>
<td>.43</td>
<td>.07</td>
</tr>
<tr>
<td>Task diversity</td>
<td>4/6/16</td>
<td>2,062</td>
<td>.29</td>
<td>.06</td>
<td>.16</td>
<td>.42</td>
<td>.33</td>
<td>.07</td>
</tr>
<tr>
<td>Job-oriented training performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Task diversity</td>
<td>1/1/2</td>
<td>184</td>
<td>.32</td>
<td>.07</td>
<td>-.59</td>
<td>1.00$^a$</td>
<td>.39</td>
<td>.09</td>
</tr>
<tr>
<td>Academic performance outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No task diversity</td>
<td>4/4/9</td>
<td>599</td>
<td>.22</td>
<td>.04</td>
<td>.12</td>
<td>.31</td>
<td>.24</td>
<td>.05</td>
</tr>
<tr>
<td>Task diversity</td>
<td>3/4/4</td>
<td>623</td>
<td>.24</td>
<td>.05</td>
<td>.14</td>
<td>.35</td>
<td>.27</td>
<td>.05</td>
</tr>
</tbody>
</table>

Note. $k$ = number of sources of data/independent study samples/effect sizes; $n$ = cumulative total sample size for effect sizes; $r_{\text{weighted}}$ = inverse sampling variance-weighted, uncorrected mean correlation; $\text{se}_r$ = standard error of weighted, uncorrected mean correlation; $r_c$ = mean correlation corrected for measurement error in the criterion; $\text{se}_{r_c}$ = standard error of correlation corrected for measurement error in the criterion; 95% CI = the lower limit (LL) and upper limit (UL) of the confidence interval around the correlation corrected for measurement error in the criterion.

$^a$ Upper limit of this confidence interval was above 1.00 but reported at the highest correlation value possible.
Table 10

Language Knowledge Requirements and Performance Outcomes

<table>
<thead>
<tr>
<th></th>
<th>$k_{sources/samples/effects}$</th>
<th>$n$</th>
<th>$r_w$</th>
<th>se_r</th>
<th>95% CI</th>
<th>$r_c$</th>
<th>se_r</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall performance outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No language knowledge requirement</td>
<td>89/103/183</td>
<td>29,084</td>
<td>.19</td>
<td>.02</td>
<td>.14</td>
<td>.23</td>
<td>.22</td>
<td>.03</td>
</tr>
<tr>
<td>Language knowledge requirement</td>
<td>19/26/51</td>
<td>7,574</td>
<td>.23</td>
<td>.05</td>
<td>.13</td>
<td>.33</td>
<td>.28</td>
<td>.05</td>
</tr>
<tr>
<td><strong>Job performance outcomes (overall)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No language knowledge requirement</td>
<td>40/47/89</td>
<td>14,005</td>
<td>.17</td>
<td>.03</td>
<td>.10</td>
<td>.23</td>
<td>.22</td>
<td>.04</td>
</tr>
<tr>
<td>Language knowledge requirement</td>
<td>9/13/33</td>
<td>5,435</td>
<td>.24</td>
<td>.08</td>
<td>.08</td>
<td>.40</td>
<td>.31</td>
<td>.09</td>
</tr>
<tr>
<td><strong>Subjective job performance outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No language knowledge requirement</td>
<td>29/34/64</td>
<td>10,707</td>
<td>.14</td>
<td>.03</td>
<td>.08</td>
<td>.20</td>
<td>.19</td>
<td>.04</td>
</tr>
<tr>
<td>Language knowledge requirement</td>
<td>3/6/11</td>
<td>2,730</td>
<td>.12</td>
<td>.11</td>
<td>-.10</td>
<td>.34</td>
<td>.17</td>
<td>.15</td>
</tr>
<tr>
<td><strong>Objective job performance outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No language knowledge requirement</td>
<td>14/15/25</td>
<td>3,298</td>
<td>.25</td>
<td>.07</td>
<td>.11</td>
<td>.38</td>
<td>.28</td>
<td>.07</td>
</tr>
<tr>
<td>Language knowledge requirement</td>
<td>7/9/22</td>
<td>22,705</td>
<td>.36</td>
<td>.10</td>
<td>.16</td>
<td>.55</td>
<td>.40</td>
<td>.11</td>
</tr>
</tbody>
</table>
Table 10 Continued

<table>
<thead>
<tr>
<th></th>
<th>$k_{\text{sources/samples/effects}}$</th>
<th>$n$</th>
<th>$r_w$</th>
<th>se$_r$</th>
<th>95% CI$_{r_w}$</th>
<th>$r_c$</th>
<th>se$_{r_c}$</th>
<th>95% CI$_{r_c}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LL</td>
<td>UL</td>
<td>LL</td>
<td>UL</td>
</tr>
<tr>
<td><strong>Job-oriented training performance outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No language knowledge requirement</td>
<td>15/17/46</td>
<td>8,445</td>
<td>.22</td>
<td>.06</td>
<td>.09</td>
<td>.34</td>
<td>.07</td>
<td>.10</td>
</tr>
<tr>
<td>Language knowledge requirement</td>
<td>1/1/2</td>
<td>184</td>
<td>.32</td>
<td>.17</td>
<td>-.02</td>
<td>.65</td>
<td>.39</td>
<td>.19</td>
</tr>
<tr>
<td><strong>Academic performance outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No language knowledge requirement</td>
<td>37/41/48</td>
<td>6,634</td>
<td>.18</td>
<td>.03</td>
<td>.14</td>
<td>.22</td>
<td>.20</td>
<td>.02</td>
</tr>
<tr>
<td>Language knowledge requirement</td>
<td>10/12/16</td>
<td>1,955</td>
<td>.20</td>
<td>.04</td>
<td>.13</td>
<td>.28</td>
<td>.22</td>
<td>.04</td>
</tr>
</tbody>
</table>

*Note. $k =$ number of sources of data/independent study samples/effect sizes; $n =$ cumulative total sample size for effect sizes; $r_{\text{weighted}} =$ inverse sampling variance-weighted, uncorrected mean correlation; se$_r =$ standard error of weighted, uncorrected mean correlation; $r_c =$ mean correlation corrected for measurement error in the criterion; se$_{r_c} =$ standard error of correlation corrected for measurement error in the criterion; 95% CI =$ the lower limit (LL) and upper limit (UL) of the confidence interval around the correlation corrected for measurement error in the criterion.
Table 11

Summary of Findings by Hypothesis

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Theoretical Rationale</th>
<th>Summary of Findings</th>
<th>Key Takeaways</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothesis 1:</strong> Information processing measures will be positively correlated with (a) job performance, (b) job-oriented training performance, and (c) academic performance.</td>
<td>Information processing entails the fundamental cognitive processes that lead to declarative and procedural knowledge, and therefore should predict performance across several domains.</td>
<td>The mean uncorrected correlation between information processing and all performance outcomes is .19.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The smallest mean uncorrected correlation is between information processing and subjective job performance at .14.</td>
<td></td>
<td>Past meta-analyses have shown other intelligence measures to have uncorrected correlations between .22 (Bertua et al., 2005) and .43 (Bobko, Roth, &amp; Potosky, 1999) with subjective job performance, .30 (Roth et al., 2005) with objective job performance, between .23 (Schmitt &amp; Fandre, 2008) and .38 (Postlethwaite, 2011) with job-oriented training, and between .27 and .45 (Kuncel &amp; Hezlett, 2007) with academic performance. While the mean uncorrected correlations for information processing measures are slightly below other measures of intelligence, they clearly have predictive validity. The hypothesis is supported.</td>
</tr>
<tr>
<td></td>
<td>The largest mean uncorrected correlation is between information processing and objective job performance at .30.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypothesis</td>
<td>Theoretical Rationale</td>
<td>Summary of Findings</td>
<td>Key Takeaways</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>---------------</td>
</tr>
<tr>
<td><strong>Hypothesis 2:</strong> Information processing measures will have smaller group scores differences between African American and Caucasian test takers than traditionally found on intelligence measures (i.e., $d = 1.00$).</td>
<td>Information processing measures reduce the reliance on prior knowledge, which can differ based on past education, cultural background, socio-economic status, and so forth. By minimizing knowledge on these measures, differences between groups should also be reduced.</td>
<td>The uncorrected $d$ value across all information processing measures is .41, with a 95% confidence interval of .17 to .65.</td>
<td>Past studies have shown other intelligence measures to have uncorrected $d$ values ranging from .60 for specific abilities (Hough et al., 2001) to 1.14 for the GATB (Berry et al., 2014). Generally, reviews point to a $d$ value of 1.00 for intelligence measures. The current study shows that information processing measures have $d$ values that are much smaller than common measures of intelligence. The hypothesis is supported.</td>
</tr>
</tbody>
</table>
Table 11 Continued

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Theoretical Rationale</th>
<th>Summary of Findings</th>
<th>Key Takeaways</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothesis 3:</strong> Cognitive science measures of information processing will have a stronger relationship with (a) job performance, (b) job-oriented training performance, and (c) academic performance than psychometric measures of information processing.</td>
<td>The two theoretical approaches differ in their theoretical foundation, coverage of the construct space, and item design. The psychometric approach relies more on statistical analyses for determining the structure of intelligence whereas the cognitive science approach uses biological and neurophysiological research. This also translates to the cognitive science approach capturing a broader set of cognitive functions as it relates directly to cognitive processes rather than relying on a statistically abstract approach. Finally, the cognitive science approach uses more dynamic stimuli compared to the psychometric approach.</td>
<td>The cognitive science approach consistently has higher uncorrected validities except for subjective performance outcomes, which slightly favors the psychometric approach with a difference of .03. For the other outcomes, the smallest differences favoring the cognitive science approach are .05 for overall performance and academic performance. The largest difference favoring the cognitive science approach is .11 for objective job performance. However, the confidence intervals for all outcomes overlap.</td>
<td>The trends generally support the notion that the cognitive science approach has stronger validities than the psychometric approach. More research is needed though since the confidence intervals overlap.</td>
</tr>
</tbody>
</table>
### Table 11 Continued

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Theoretical Rationale</th>
<th>Summary of Findings</th>
<th>Key Takeaways</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothesis 4:</strong> Cognitive science measures of information processing will have smaller group score differences between African American and Caucasian test takers than psychometric measures of information processing.</td>
<td>The cognitive science approach is more deeply rooted in the physiological sciences than the psychometric approach. Research suggests that group score differences cannot be fully explained by biological factors, meaning measures assessing the physiological domain may reduce deficiencies and contamination that are associated with irrelevant error contributing to group score differences.</td>
<td>The uncorrected $d$ value for the cognitive science approach (.40) is lower than the uncorrected $d$ value for the psychometric approach (.46). However, the confidence intervals overlap.</td>
<td>Both approaches lead to smaller differences than seen with typical intelligence measures. There is a trend suggesting the cognitive science approach produces smaller group score differences than the psychometric approach, though the confidence intervals overlap.</td>
</tr>
</tbody>
</table>
Table 11 Continued

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Theoretical Rationale</th>
<th>Summary of Findings</th>
<th>Key Takeaways</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothesis 5:</strong> Psychometric measures of information processing with a diverse set of item types will have a stronger relationship with (a) job performance, (b) job-oriented training performance, and (c) academic performance than psychometric measures of information processing that do not have a diverse set of item types.</td>
<td>Increasing the number of tasks assessed within a measure should reduce error associated with any one of those measures. Therefore, more tasks within a measure should lower the psychometric sampling error of the measure.</td>
<td>For most outcomes, there is either little difference between the uncorrected effect sizes or the uncorrected effect size for measures with no task diversity is larger than for measures with task diversity. Job-oriented training has the largest difference favoring no task diversity with a difference of .19, though the sample size is very small. The next largest differences favoring no task diversity are .08 for all performance outcomes and academic performance. There is a small difference (.02) favoring measures with task diversity for subjective job performance. The confidence intervals overlap for all outcomes.</td>
<td>Overall, the hypothesis is not supported as the trends mostly suggest larger validities for measures with no task diversity than measures with task diversity. The small sample sizes and large overlap in the confidence intervals means more research is needed to understand this pattern of findings.</td>
</tr>
<tr>
<td>Hypothesis</td>
<td>Theoretical Rationale</td>
<td>Summary of Findings</td>
<td>Key Takeaways</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>---------------</td>
</tr>
<tr>
<td><strong>Hypothesis 6:</strong> Cognitive science measures of information processing with a diverse set of item types will have a stronger relationship with (a) job performance, (b) job-oriented training performance, and (c) academic performance than cognitive science measures of information processing that do not have a diverse set of item types.</td>
<td>Measures from the cognitive science approach are meant to assess specific cognitive functions. The more tasks included in a cognitive science measure should increase the number of specific functions being assessed, and therefore reduce deficiencies in assessing the focal variable.</td>
<td>For most outcomes, measures with task diversity have a larger mean uncorrected correlation than measures with no task diversity. The largest differences favoring measures with task diversity are .12 for subjective job performance and academic performance, though the sample sizes are small. There is a difference favoring measures with no task diversity for objective job performance (.09). Overall, the confidence intervals overlap for all outcomes.</td>
<td>The trends generally support the hypothesis. Measures with task diversity typically have larger validities than measures with no task diversity. However, there is a limited sample size and the confidence intervals overlap. Also, the validity for objective job performance contradicts the hypothesis. More research is needed to better understand the trends related to this hypothesis.</td>
</tr>
</tbody>
</table>
Table 11 Continued

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Theoretical Rationale</th>
<th>Summary of Findings</th>
<th>Key Takeaways</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothesis 7:</strong> Information processing measures that require less language knowledge will have a stronger relationship with (a) job performance, (b) job-oriented training performance, and (c) academic performance than information processing measures that require more language knowledge.</td>
<td>Knowledge language requirements can introduce irrelevant test content when the performance outcome does not require the language assessed on the measure. Therefore, reducing the language knowledge requirement should reduce potential contamination from irrelevant knowledge.</td>
<td>For most outcomes, measures with a language knowledge requirement have a larger mean uncorrected correlation than measures with no language knowledge requirement. The largest difference favoring measures with a language knowledge requirement is .11 for objective job performance. There is a small difference (.02) favoring measures with no language knowledge requirements. Overall, the confidence intervals overlap for all outcomes.</td>
<td>The trends oppose the original hypothesis as most outcomes have a higher validity for measures that require language knowledge than measures that do not require language knowledge. The confidence intervals overlap across for all outcomes, meaning more research is needed to better understand these trends.</td>
</tr>
</tbody>
</table>
Figure 1. Decision tree for inclusion/exclusion criteria of meta-analysis.
References

* References denoted with an asterisk were included in the validity meta-analysis.
† References denoted with an obelisk were included in the group score differences meta-analysis.


Cattell, R. B., & Bristol, H. (1933). Intelligence tests for mental ages of four to eight years. *British Journal of Educational Psychology, 3,* 142-169.


Lievens, F., & Reeve, C. L. (2012). Where I-O psychology should really (re)start its investigation of intelligence constructs and their measurement. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 5*(2), 153-158.


Oswald, F. L., & Hough, L. (2012). I–O 2.0 from intelligence 1.5: Staying (just) behind the cutting edge of intelligence theories. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 5*, 172-175.


